

UNIVERSIDADE FEDERAL DO PARANÁ

JULIO CESAR WOJCIECHOWSKI

ESTIMATIVA DE VOLUME COM MODELOS GEOESTATÍSTICOS DE  
PARÂMETROS COMPARTILHADOS



CURITIBA

2014

JULIO CESAR WOJCIECHOWSKI

ESTIMATIVA DE VOLUME COM MODELOS GEOESTATÍSTICOS DE  
PARÂMETROS COMPARTILHADOS

Tese apresentada ao Programa de Pós-graduação em Engenharia Florestal, Área de Concentração em Manejo Florestal, Departamento de Ciências Florestais, Setor de Ciências Agrárias, Universidade Federal do Paraná, como parte das exigências para a obtenção do título de Doutor em Engenharia Florestal.

Orientador: Prof. Dr. Julio Eduardo Arce

Co-orientadores: Prof. Dr. Saulo Henrique Weber e Prof.  
Dr. Carlos Alberto da Fonseca Pires

CURITIBA

2014

Ficha catalográfica elaborada por Denis Uezu – CRB 1720/PR  
Biblioteca de Ciências Florestais e da Madeira - UFPR

Wojciechowski, Julio Cesar

Estimativa de volume com modelos geoestatísticos de parâmetros  
compartilhados / Julio Cesar Wojciechowski. – 2014  
113 f. : il.

Orientador: Prof. Dr. Julio Eduardo Arce

Coorientadores: Prof. Dr. Saulo Henrique Weber

Prof. Dr. Carlos Alberto da Fonseca Pires

Tese (doutorado) - Universidade Federal do Paraná, Setor de Ciências  
Agrárias, Programa de Pós-Graduação em Engenharia Florestal. Defesa:  
Curitiba, 12/08/2014.

Área de concentração: Manejo Florestal

1. Levantamentos florestais – Métodos estatísticos. 2. Dendrometria. 3.  
Teses. I. Arce, Julio Eduardo. II. Weber, Saulo Henrique. III. Pires, Carlos Alberto  
da Fonseca. IV. Universidade Federal do Paraná, Setor de Ciências Agrárias. V.  
Título.

CDD – 634.9

CDU – 634.0.524.6




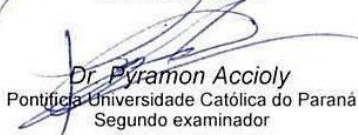
Universidade Federal do Paraná  
Setor de Ciências Agrárias - Centro de Ciências Florestais e da Madeira  
Programa de Pós-Graduação em Engenharia Florestal

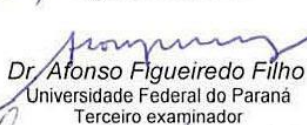
## PARECER

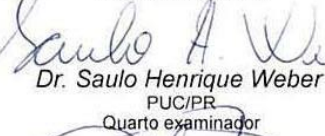
Defesa nº. 1050

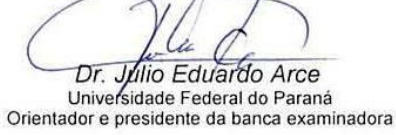
A banca examinadora, instituída pelo colegiado do Programa de Pós-Graduação em Engenharia Florestal, do Setor de Ciências Agrárias, da Universidade Federal do Paraná, após arguir o(a) doutorando(a) *Julio Cesar Wojciechowski* em relação ao seu trabalho de tese intitulado **"ESTIMATIVA DE VOLUME COM MODELOS GEOESTATÍSTICOS DE PARÂMETROS COMPARTILHADOS"**, é de parecer favorável à APROVAÇÃO do(a) acadêmico(a), habilitando-o(a) ao título de *Doutor* em Engenharia Florestal, área de concentração em MANEJO FLORESTAL.

  
Dr. Andrea Nogueira Dias  
Universidade Estadual do Centro-Oeste  
Primeiro examinador

  
Dr. Pyramon Accioly  
Pontifícia Universidade Católica do Paraná  
Segundo examinador

  
Dr. Afonso Figueiredo Filho  
Universidade Federal do Paraná  
Terceiro examinador

  
Dr. Saulo Henrique Weber  
PUC/PR  
Quarto examinador

  
Dr. Julio Eduardo Arce  
Universidade Federal do Paraná  
Orientador e presidente da banca examinadora



Curitiba, 12 de agosto de 2014.

  
Antonio Carlos Batista  
Coordenador do Curso de Pós-Graduação em Engenharia Florestal

À minha esposa, filho e mãe, que sempre foram meu maior motivo para seguir em frente, aos meus queridos mestres e amigos.

DEDICO

## **AGRADECIMENTOS**

À Universidade Federal do Paraná e a Pós-graduação em Engenharia Florestal, pela oportunidade de fazer algo tão importante na minha vida profissional.

Ao CNPq pela concessão da bolsa de doutorado.

Ao meu orientador prof. Dr. Julio Eduardo Arce pelo crédito ao meu trabalho, conhecimentos transmitidos, compreensão nos momentos difíceis e pelo modo atencioso e bem humorado no qual sempre foram conduzidas nossas conversas.

À Universidade do Estado do Mato Grosso e ao quadro de professores do Departamento de Engenharia Florestal, em especial aos amigos e colegas de trabalho, Everton Valdomiro Pedroso Brun e Edgley Pereira da Silva.

Ao prof. Dr. Paulo Justiniano Ribeiro Jr. pelo apoio, disponibilidade e visão científica que deram novas perspectivas ao trabalho. A Naimara Vieira do Prado pela ajuda na construção dos códigos em R que foram base do estudo.

Aos meus Co-orientadores, prof. Dr. Saulo Henrique Weber pelos ensinamentos e disponibilidade incondicional em ajudar, e prof. Dr. Carlos Alberto da Fonseca Pires pela amizade, atenção e apoio que sempre me prestou desde o mestrado.

Aos professores do Curso de Pós-graduação da UFPR, Dr. Franklin Galvão, Dr. Carlos Vellozo Roderjan, Dr. Sylvio Péllico Neto, Dr. Sebastião do Amaral Machado, Dr. Nelson Yoshihiro Nakajima e Dr<sup>a</sup>. Yoshiko; mestres e amigos que sempre me ofereceram sincera atenção e conhecimento.

Aos meus amigos e colegas da pós-graduação, que sempre lembrarei com muito carinho, em especial a Tomaz Longhi; Jaçanan Eloisa Milani, Uberson Boaretto Rossa, Thiago F. Stepka, Rodrigo Lima, Gerson Lisboa e Flavio Nascimento.

À secretaria da Pós-graduação em nome do Reinaldo Mendes de Souza e David Teixeira de Araújo, pelo modo prestativo e atencioso que sempre fui tratado desde o primeiro dia no curso.

À família, em especial à minha companheira e esposa Luzia Venturini, meu filho Anael de S. Wojciechowski e minha mãe Thereza M. de A. Wojciechowski, obrigado por fazerem parte da minha vida.

## **BIOGRAFIA**

JULIO CESAR WOJCIECHOWSKI, filho de Thereza Marina de A. Wojciechowski nasceu em Santa Maria no Rio Grande do Sul, no dia 02 de Agosto de 1974.

Em 1997 iniciou o curso de Engenharia Florestal pela Universidade Federal de Santa Maria (UFSM), no decorrer do curso participou de projetos de pesquisa e extensão dentre os quais destacam-se: Dinâmica De Três Fases Sucessionais De Uma Floresta Estacional Decidual Do RS, Inventário Florestal Contínuo do RS, Inventário Florestal da vegetação da área de alagamento em SC e "Floresta de Araucária e suas transições" Peld-CNPq pela UFPR.

Formou-se Engenheiro Florestal no ano de 2003, no mesmo ano atuou como profissional prestando serviços de auditoria ambiental na empresa Rio Grande Energia (RGE). Em 2004 iniciou o curso de Mestrado na área de concentração "Tecnologia da Geoinformação", obtendo o título de Mestre em Geomática pela Universidade Federal de Santa Maria (UFSM) no ano de 2006.

Nesse mesmo ano ingressou no quadro efetivo do Departamento de Engenharia Florestal pela Universidade do Estado do Mato Grosso (UNEMAT) onde é professor na área de Manejo Florestal, ministrando disciplinas de Manejo Florestal, Dendrometria, Inventário Florestal, Sensoriamento Remoto e Geoprocessamento para fins florestais.

Em 2009 foi convidado a participar do livro "A FLORESTA ESTACIONAL SUBTROPICAL - Caracterização e ecologia no rebordo do Planalto Meridional" tendo dois capítulos de sua autoria na compilação.

Ingressou em 2010 no curso de Pós-Graduação em nível de Doutorado pela Universidade Federal do Paraná (UFPR), onde em 2014, obteve o título de Doutor em Engenharia Florestal - Área de Concentração: Manejo Florestal.

“A jornada é a recompensa”

Steve Paul Jobs



## RESUMO

O presente estudo teve como objetivo utilizar o compartilhamento de parâmetros do modelo geoestatístico aplicado aos estimadores de máxima verossimilhança para prever os volumes por hectare em três fragmentos de Floresta Estacional Decidua empregando a abordagem “*Borrow Strength*”. Os fragmentos separados geograficamente possuem idades pós-intervenção de 30, 55 e 75 anos e localizam-se no Município de Santa Tereza-RS. A abordagem do estudo consistiu na aplicação de métodos de estimativa conjunta, em que áreas consideradas semelhantes *a priori* para uma determinada variável compartilham informação dentro de um modelo geoestatístico único, melhorando as estimativas e suprimindo diferenças de intensidade amostral entre os fragmentos. Os dados foram coletados em 56 unidades amostrais (U.A) de tamanho variável com aproximadamente 250 m<sup>2</sup> num total de nove ha, distribuídas em um *grid* sistemático de 40 x 40 m, sendo medidas as variáveis dendrométricas dos indivíduos com DAP  $\geq 10$  cm próximas ao centro das unidades. Foram elaboradas duas abordagens para o conjunto de dados, sendo que a primeira considerou as áreas totalmente independentes entre si, subdivididas em dois tipos: ajuste ao modelo não espacial (NSM) e ajuste pelo método de máxima verossimilhança (MV) não compartilhado (ajuste individual). A segunda abordagem descreveu os ajustes dos modelos compartilhados em função do erro aleatório ou *nugget*, sendo: modelos compartilhados sem *nugget* fixo (variabilidade entre as U.A) e modelos compartilhados com *nugget* fixo (variabilidade dentro das U.A). Todos os modelos com exceção dos modelos NSM, foram ajustados utilizando a função de máxima verossimilhança logarítmica aplicada ao modelo de correlação exponencial da família Matérn. Em seguida os modelos foram comparados pelo critério de informação de Akaike (AIC) e também pelo grau de dependência espacial para posterior krigagem e elaboração das superfícies de predição dos modelos selecionados. Foi observado que os modelos combinados para estimativa do volume foram superiores para os valores de AIC e grau de dependência espacial em relação aos ajustes para as áreas individuais. Entre os modelos compartilhados, observou-se que houve um ganho nas estimativas dos parâmetros utilizando o *nugget* fixo, que resultaram em uma correlação das amostras e grau de dependência espacial maior (AP = 88 m), em relação aos modelos compartilhado sem *nugget* fixo (AP = 75 e 66 m). O valor do AIC mostrou-se eficiente uma vez que comparou os diferentes níveis de ajustes propostos na metodologia do trabalho, selecionando um modelo com parcimônia e compatível com os padrões de distribuição espacial encontrados nas áreas. As estimativas de superfície retornaram a variável aos seus valores originais com boa precisão. Sugere-se o uso de modelos combinados para dados de amostragem em diferentes áreas, com introdução da estimativa do erro intra-parcela (*nugget* fixo) nas equações de MV, para aumentar a correlação entre as U.A, e avaliação conjunta do AIC somado ao grau de dependência espacial na estimativa de variáveis dendrométricas.

Palavras-chave: Máxima verossimilhança; *borrow strength*; modelos combinados; Inventário Florestal

## ABSTRACT

This study aimed to use the share parameters of the geostatistical model applied to maximum likelihood estimators to predict the volumes per hectare in three fragments of Deciduous Forest employing the "Borrow Strength" approach. The geographically separated fragments have age's post-intervention 30, 55 and 75 years and are located in the municipality of Santa Teresa-RS. The approach of the study was the application of methods for combined estimates, where similar areas considered a priori for a given variable share information within a single geostatistical model, improving the estimates and supplying differences in sampling intensity among fragments. Data were collected in 56 sampling units (S.U.) of variable sizes with approximately 250 m<sup>2</sup> for a total of nine ha, distributed in a systematic grid of 40 x 40 m. Dendrometric variables from individuals with DBH  $\geq$  10 cm near the center of the S.U. were measured. Two approaches to the data set were prepared, the first of which considering both areas entirely independent themselves, subdivided into two types: a fit to non-spatial model (NSM) and a fit to the maximum likelihood (ML) not shared (individual adjustment) model. The second approach described the adjustment of the shared as a function of random error or nugget, comprising models: a shared model without fixed nugget (variability between S.U.) and a shared model with fixed nugget (variability within S.U.) models. All models except the NSM were fitted using a logarithmic function of maximum likelihood applied to the *Matérn* family of exponential correlation model. Then, the models were compared using Akaike information criterion (AIC) and also by the degree of spatial dependence for subsequent preparation of both kriging and prediction surfaces of the selected models. It was observed that the combined volume models to estimate values were higher for the AIC values and spatial dependence with respect to the adjustments for the individual areas. Among the shared models, it was observed that there was a gain in the parameter estimates using the fixed nugget, which resulted in a higher correlation of samples and spatial dependence (AP = 88 m), than the shared models without the fixed nugget (AP = 75 and 66 m). The AIC value was efficient because it compared the different levels of proposed adjustments to the methodology of the study, selecting a model with parsimony and compatible with the spatial distribution patterns found in the areas. Estimates of the surface returned the variable to its original values with improved accuracy. The use of combined models for data sampling in different areas with the introduction of the error estimate intra-plot (fixed nugget) in the equations of MV can be suggested to increase the correlation between the S.U. and combined evaluation of the AIC plus the degree of spatial dependence in estimating dendrometric variables.

Key-words: Maximum likelihood; borrow strength; combined models; Forest Inventory

## LISTA DE FIGURAS

FIGURA 1 – GRÁFICO DO SEMIVARIOGRAMA EXPERIMENTAL COM OS PARÂMETROS EFEITO PEPITA ( $\tau^2$ ), VARIÂNCIA ( $\sigma^2$ ) E ALCANÇE ( $\Phi$ ).....	26
FIGURA 2 – AMOSTRAS COM ESPAÇAMENTO REGULAR DE 40 X 40 METROS, PARA O CÁLCULO DO VALOR DE “Z” .....	33
FIGURA 3 - PARÂMETROS PARA O CÁLCULO DO SEMIVARIOGRAMA COM BASE EM AMOSTRAS IRREGULARMENTE ESPAÇADAS EM DUAS DIMENSÕES.....	34
FIGURA 4 – LOCALIZAÇÃO DOS FRAGMENTOS: (FEI) FLORESTA EM ESTÁGIO INICIAL, (FES) FLORESTA EM ESTÁGIO SECUNDÁRIO E (FMA) FLORESTA MADURA.....	39
FIGURA 5 - DETALHE DO MICRO RELEVO NAS ÁREAS EXPERIMENTAIS: A) FEI, B) FES E C) FMA.....	40
FIGURA 6 - FASES DA COLETA DE SOLO A) ABERTURA DA TRINCHEIRA; B) MEDIÇÃO DA PROFUNDIDADE; C) COLETA DAS ANÁLISES QUÍMICAS E D) COLETA DAS ANÁLISES FÍSICAS.....	41
FIGURA 7 - ESQUEMA DE AMOSTRAGEM UTILIZADO NOS FRAGMENTOS: A) FEI; B) FES E C) FMA; D) DETALHE DA UNIDADE AMOSTRAL.....	44
FIGURA 8 – FLUXOGRAMA DAS ETAPAS DA METODOLOGIA DE ANÁLISE E TRATAMENTO DOS DADOS.....	56
FIGURA 9 - GRÁFICOS <i>BOX PLOT</i> DA DISPERSÃO DOS VOLUMES ( $\text{m}^3 \text{ha}^{-1}$ ) NAS ÁREAS DE A) FEI; B) FES E C) FMA.....	58
FIGURA 10 – A) PARCELAS SUBDIVIDIDAS EM <i>QUANTIS</i> , DISPERSÃO DOS VALORES CONTRA AS COORDENADAS B) X E C) Y, D) HISTOGRAMA (VOL/HA) NA ÁREA DE FEI.....	59
FIGURA 11 – A) PARCELAS SUBDIVIDIDAS EM <i>QUANTIS</i> , DISPERSÃO DOS VALORES CONTRA AS COORDENADAS B) X E C) Y, D) HISTOGRAMA (VOL/HA) NA ÁREA DE FES.....	60
FIGURA 12 – A) PARCELAS SUBDIVIDIDAS EM <i>QUANTIS</i> , DISPERSÃO DOS VALORES CONTRA AS COORDENADAS B) X E C) Y, D) HISTOGRAMA (VOL/HA) NA ÁREA DE FMA.....	60

FIGURA 13 - PERFIL DA LOG-VEROSSIMILHANÇA PARA O PARÂMETRO DE TRANSFORMAÇÃO BOX-COX APLICADO A VARIÁVEL VOLUME NAS ÁREAS: A) FEI, B) FES E C) FMA.....	61
FIGURA 14 - GRÁFICOS QUANTIL-QUANTIL (Q-Q) DA DISTRIBUIÇÃO NORMAL PARA O VOLUME NAS ÁREAS EXPERIMENTAIS: A) FEI, B) FES E C) FMA.....	62
FIGURA 15 – PREDIÇÃO DOS VOLUMES ( $\text{m}^3 \text{ ha}^{-1}$ ) PARA O MODELO COMBINADO SEM <i>NUGGET</i> FIXO “04” NAS ÁREAS DE FEI, FES E FMA.....	82
FIGURA 16 – PREDIÇÃO DOS VOLUMES ( $\text{m}^3 \text{ ha}^{-1}$ ) PARA O MODELO COMPARTILHADO COM <i>NUGGET</i> FIXO “11” NAS ÁREAS DE FEI, FES E FMA.....	83
FIGURA 17 – PREDIÇÃO DOS VOLUMES ( $\text{m}^3 \text{ ha}^{-1}$ ) PARA O MODELO NÃO COMPARTILHADO DE MÁXIMA VEROSSIMILHANÇA NAS ÁREAS DE FEI, FES E FMA.....	84

## LISTA DE TABELAS

TABELA 1 - CODIFICAÇÃO DAS POSIÇÕES 111, 112, 121, 122 E 123 EM FUNÇÃO DOS CRUZAMENTOS DAS ÁREAS NA SEQUÊNCIA FEI, FES, FMA.....	52
TABELA 2 - ESTIMATIVA DE MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME PARA OS MODELOS DE PARÂMETROS NÃO COMPARTILHADOS NAS ÁREAS FEI-FES-FMA.....	63
TABELA 3 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ , $\tau^2$ , $\sigma^2$ e $\Phi$ ) DO MODELO NAS ÁREAS FEI-FES-FMA (MODELO MÍNIMO).....	65
TABELA 4 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\sigma^2$ e $\Phi$ ) DO MODELO NAS ÁREAS FEI-FES-FMA.....	65
TABELA 5 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ , $\tau^2$ , $\sigma^2$ e $\Phi$ ) DO MODELO NAS ÁREAS FEI-FES.....	67
TABELA 6 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ e $\tau^2$ ) DO MODELO NAS ÁREAS FEI-FES E ( $\Phi$ ) NAS ÁREAS FES-FMA.....	68
TABELA 7 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ ) DO MODELO NAS ÁREAS FEI-FES E ( $\sigma^2$ e $\Phi$ ) NAS ÁREAS FES-FMA.....	69
TABELA 8 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ , $\tau^2$ e $\Phi$ ) DO MODELO NAS ÁREAS DE FEI-FES.....	70
TABELA 9 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO O PARÂMETRO ( $\sigma^2$ ) DO MODELO NAS ÁREAS FEI-FES-FMA.....	70
TABELA 10 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COM PARÂMETROS ( $\beta$ , $\tau^2$ , $\sigma^2$ e $\Phi$ ) LIVRES NAS ÁREAS FEI-FES-FMA (MODELO MÁXIMO).....	71
TABELA 11 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ , $\sigma^2$ e $\Phi$ ) DO MODELO NAS ÁREAS FEI-FES-FMA (MODELO MÍNIMO-NUGGET FIXO).....	72

TABELA 12 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ e $\sigma^2$ ) DO MODELO NAS ÁREAS FEI-FES-FMA E ( $\Phi$ ) NAS ÁREAS FEI-FES (NUGGET FIXO).....	73
TABELA 13 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ E $\Phi$ ) DO MODELO NAS ÁREAS FEI-FES E ( $\sigma^2$ ) NAS ÁREAS FES-FMA (NUGGET FIXO).....	73
TABELA 14 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO O PARÂMETROS ( $\sigma^2$ ) DO MODELO NAS ÁREAS FES-FMA E ( $\Phi$ ) NAS ÁREAS FEI-FMA (NUGGET FIXO).....	74
TABELA 15 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COM OS PARÂMETROS ( $\beta$ , $\sigma^2$ e $\Phi$ ) DO MODELO NAS ÁREAS FEI-FES-FMA (MODELO MÁXIMO-NUGGET FIXO).....	75
TABELA 16 - SELEÇÃO DOS MODELOS COM AJUSTE LINEAR SEM COMPONENTE ESPACIAL (NSM) E POR MÁXIMA VEROSSIMILHANÇA (MV) NAS ÁREAS FEI-FES-FMA PELO CRITÉRIO DE AKAIKE (AIC).....	76
TABELA 17 - SELEÇÃO DOS MODELOS COM AJUSTE POR MÉTODO DE COMPARTILHAMENTO DE PARÂMETROS NAS ÁREAS FEI, FES, FMA PELO CRITÉRIO DE INFORMAÇÃO DE AKAIKE (AIC).....	78
TABELA 18 - SIMILARIDADE DOS MODELOS DE COMPARTILHAMENTO DE PARÂMETROS COM MAIOR GRAU DE DEPENDÊNCIA ESPACIAL ( $\Phi$ ) NAS ÁREAS FEI, FES, FMA.....	79
TABELA 19 - SEPARABILIDADE DOS PARÂMETROS DO MODELO “11” EM RELAÇÃO AO GRAU DE DEPENDÊNCIA ESPACIAL ( $\Phi$ ).....	80
TABELA 20 - SEPARABILIDADE DOS PARÂMETROS DO MODELO “06” EM RELAÇÃO AO GRAU DE DEPENDÊNCIA ESPACIAL ( $\Phi$ ).....	80
TABELA 21 – SEPARABILIDADE DOS PARÂMETROS DO MODELO “04” EM RELAÇÃO AO GRAU DE DEPENDÊNCIA ESPACIAL ( $\Phi$ ).....	81
TABELA 22 – RESULTADO DA VALIDAÇÃO CRUZADA PARA OS DADOS DE VOLUME UTILIZANDO O MODELO “11” NAS ÁREAS FEI E FES POR UNIDADE AMOSTRAL.....	85
TABELA 23 – COMPARAÇÃO ENTRE O ESTIMADOR CLÁSSICO (PRODAN) E O ESTIMADOR GEOESTATÍSTICO (KRIGAGEM) PARA O VOLUME TOTAL.....	86

## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>17</b>
<b>2 OBJETIVOS.....</b>	<b>20</b>
2.1 OBJETIVO GERAL.....	20
2.2 OBJETIVOS ESPECÍFICOS.....	20
<b>3 REFERENCIAL TEÓRICO.....</b>	<b>21</b>
3.1 FLORESTA ESTACIONAL DECIDUAL .....	21
3.2 ANÁLISE EXPLORATÓRIA DE DADOS (A.E.D) .....	22
3.3 VARIÁVEIS REGIONALIZADAS .....	23
3.4 HIPÓTESES ESTACIONÁRIAS .....	24
3.5 MÉTODOS DE ESTIMATIVAS EM GEOESTATÍSTICA .....	24
3.5.1 Métodos de estimativa por semivariograma .....	25
3.5.2 Métodos de estimativa por máxima verossimilhança .....	28
3.6 CRITÉRIO DE INFORMAÇÃO DE AKAIKE (AIC) .....	29
3.7 MODELOS MISTOS E MODELOS MISTOS COMPARTILHADOS .....	30
3.8 PLANO DE AMOSTRAGEM PARA ANÁLISE DA DEPENDÊNCIA ESPACIAL .....	31
3.8.1 Levantamentos com <i>grid</i> regular .....	32
3.8.2 Levantamentos com <i>grid</i> irregular .....	33
3.9 KRIGAGEM E PREDIÇÃO.....	35
3.10 APLICAÇÕES DA GEOESTATÍSTICA NOS LEVANTAMENTOS FLORESTAIS .....	37
<b>4 MATERIAL E MÉTODOS .....</b>	<b>39</b>
4.1 CARACTERIZAÇÃO DA ÁREA DE ESTUDO.....	39
4.1.1 Aspectos gerais de clima, solos e relevo.....	40
4.1.2 Aspectos fitofisionômicos .....	42
4.2 OBTENÇÃO DOS DADOS .....	43
4.3 ANÁLISE EXPLORATÓRIA DE DADOS.....	46
4.4 DEPENDÊNCIA ESPACIAL DA VARIÁVEL VOLUME ESTIMADA POR VEROSSIMILHANÇA.....	47
4.4.1 Ajuste de modelos sem componente espacial .....	48
4.4.2 Ajuste de modelos espaciais não compartilhados .....	49

4.4.3 Ajuste de modelos com compartilhamento de força .....	50
4.4.4 Construção de modelos compartilhados .....	51
4.5 CRITÉRIOS DE SELEÇÃO DOS MODELOS PARA ESTIMATIVA DE VOLUME ....	
.....	53
4.6 KRIGAGEM E PREDIÇÃO.....	54
4.7 VALIDAÇÃO CRUZADA.....	54
4.8 RESUMO DAS ETAPAS METODOLÓGICAS .....	55
<b>5 RESULTADO E DISCUSSÃO .....</b>	<b>57</b>
5.1 ESTATÍSTICA DESCRITIVA E ANÁLISE EXPLORATÓRIA.....	57
5.2 AJUSTES DE MODELOS GEOESTATÍSTICOS POR MÁXIMA	
VEROSSIMILHANÇA PARA ESTIMATIVA DO VOLUME.....	63
5.2.1 Ajuste de modelos não compartilhados.....	63
5.2.2 Ajuste de modelos compartilhados.....	64
5.2.2.1 Ajuste de modelos compartilhados sem <i>nugget</i> fixo .....	64
5.2.2.2 Ajuste de modelos compartilhados com <i>nugget</i> fixo .....	71
5.3 SELEÇÃO DO MODELO PELO CRITÉRIO DE AKAIKE (AIC). ....	75
5.3.1 Modelos não compartilhados para estimativa do volume por hectare .....	75
5.3.2 Modelos compartilhados para estimativa do volume por hectare .....	76
5.4 KRIGAGEM E PREDIÇÃO DO VOLUME POR HECTARE .....	81
5.4.1 Krigagem e predição do modelo compartilhado sem <i>nugget</i> fixo .....	82
5.4.2 Krigagem e predição do modelo compartilhado com <i>nugget</i> fixo .....	83
5.4.3 Krigagem e predição do modelo não compartilhado .....	84
5.5 VALIDAÇÃO CRUZADA .....	84
<b>6 CONCLUSÕES .....</b>	<b>87</b>
<b>7 RECOMENDAÇÕES.....</b>	<b>89</b>
<b>REFERÊNCIAS.....</b>	<b>90</b>
<b>ANEXOS (CÓDIGOS EM R).....</b>	<b>98</b>



## 1 INTRODUÇÃO

Em estudos de inventário e mensuração florestal, a obtenção de algumas características como volume, área basal, altura da árvore, altura dominante, biomassa, entre outras, dependendo do tamanho da população a ser mensurada, ou ainda, dificuldades inerentes à logística nos locais e acesso, podem resultar em horas de trabalho e elevados custos operacionais, mesmo em processos de amostragem.

Os custos relacionados à coleta de informações de variáveis dendrométricas em florestas naturais aumentam proporcionalmente com o nível de precisão exigido, principalmente nos levantamentos florestais de grandes extensões, onde os planos de amostragem necessitam de maior complexidade e envolvem outras áreas do conhecimento que fazem uso, por exemplo, de geotecnologias associadas nas suas fases de elaboração, planejamento e execução.

Devido aos custos, entre outros fatores, muitas vezes na modelagem quantitativa ajustam-se modelos aos dados coletados em uma determinada área, sendo os modelos aceitos muitas vezes testados, adaptados e aplicados a outros locais ou regiões em condições semelhantes. Os resultados são comparados e então, é escolhida a melhor estratégia a ser adotada para solução de um problema proposto (MATTHIOPOULOS, 2011).

Uma das variáveis importantes e básicas a ser respondida em um levantamento ou inventário florestal é sem dúvida a quantificação do estoque, ou o volume da floresta. Machado e Figueiredo Filho (2003) comentaram que estimar o volume das árvores se configura como umas das principais finalidades dos levantamentos florestais.

No contexto da amostragem para subpopulações em fragmentos florestais, diversas técnicas de modelagem avançadas são aplicadas para resolver problemas relacionados às estimativas e previsões dos volumes e demais variáveis dendrométricas importantes. Técnicas baseadas na utilização de modelos ditos mistos “*mix-models*” (PINHEIRO; BATES, 2000), modelos com regressão condicional (BAS; VAN PUIJENBROEK, 2010), modelos hierárquicos (FINLEY, 2010), modelos heurísticos e meta-heurísticos (RONGXIA; PETE; KEVIN, 2010),

modelos de multicritério (MALCZEWSKI, 1999; BUONGIORNO; GILLESS, 2003) e muitos outros aplicados as diferentes áreas do conhecimento, são cada vez mais necessários no ramo da engenharia e pesquisa florestal, em resposta ao crescente volume de dados manipulados e a complexidade das informações advindas de diferentes fontes de levantamentos.

A importância em estimar com precisão não apenas uma variável em suas medidas usuais, mas também como ela distribui-se no espaço, abre caminho para inclusão de abordagens geoestatísticas criando novas concepções, já que na amostragem clássica e análise dos dados da maioria dos inventários e levantamentos florestais, a estrutura espacial e sua correlação são na maioria das vezes desconsideradas (OLIVEIRA, 1991).

Ocorre que com exceção dos modelos geoestatísticos, os modelos empregados não consideram a posição da variável e sua configuração no espaço, sendo essa uma informação importante quando se deseja obter estimativas pontuais de maneira rápida, confiável e com menor número de amostras possíveis tomadas a campo.

Atualmente é consenso que determinadas variáveis distribuem-se de maneira estruturada no espaço, algumas apresentando uma distribuição bem definida, enquanto outras com uma aparente correlação, porém, com um padrão menos estruturado. Scolforo e Mello (2006) comentaram que havendo estrutura de dependência espacial, ou seja, correlação entre os pontos amostrados, deve-se fazer uso dos benefícios de tal correlação com a introdução do componente espacial nos modelos, pois, nesse caso a aplicação da teoria clássica de amostragem em levantamentos florestais pode aumentar os problemas estatísticos, já que essa teoria baseia-se nos princípios de independência entre as observações, sendo completamente oposta ao aplicado na geoestatística, que correlaciona o grau de dependência de uma variável em relação a ela mesma no espaço, bem como, mede a força em que ocorre essa correlação.

Como resultado da aplicação da abordagem clássica para o cálculo dos estimadores em locais com estrutura espacial detectável, cada área pode erroneamente receber o mesmo tipo de tratamento ou número de amostras, dependente apenas, da variabilidade de suas médias e demais quantidades estatísticas para auxiliar na escolha do plano de amostragem adotado, ou tipo de estratificação da vegetação.

Áreas com essas características, quando fragmentadas, mesmo em tipologias semelhantes são tratadas do ponto de vista estatístico separadamente, ou seja, cada área contribui com seus valores de volumes médios e demais variáveis dendrométricas usualmente coletadas.

Nesse sentido, em um inventário baseado em várias subpopulações com classes consideradas semelhantes *à priori*, podem ser empregadas técnicas de “*Borrow Strenght*” para emprestar “força”, no sentido da informação de uma área para outra, com a finalidade de explicar o comportamento espacial de determinada variável utilizando modelos compartilhados ou combinados, bem como, ampliar as possibilidades de análises e cruzamentos advindos destas combinações (KANGAS; MALTANO, 2006). Dentro de uma concepção em que áreas semelhantes nas suas fitofisionomias, mesmo que com variações distintas de solo, relevo, idade, diferentes escalas de amostragem e tamanho de amostras podem, em uma estimativa conjunta, emprestar informação daquilo em que lhes é comum por meio de uma distribuição conhecida.

Da mesma forma, a partir desta abordagem, é possível generalizar dados de volume e outras variáveis dendrométricas obtidas na amostra, cobrindo as árvores para as quais apenas as medidas básicas estão disponíveis, como o diâmetro, obtendo resultados com potencial de unir informações de caráter dendrométrico e estimativas de caráter espacial.

## 2 OBJETIVOS

### 2.1 OBJETIVO GERAL

Utilizar o compartilhamento dos parâmetros do modelo geoestatístico aplicado aos estimadores de máxima verossimilhança para predizer os volumes por hectare em três fragmentos de Floresta Estacional Decidual empregando uma abordagem “*Borrow Strenght*”.

### 2.2 OBJETIVOS ESPECÍFICOS

- a) Ajustar modelos geoestatísticos pelo método da máxima verossimilhança (MV) aplicado às áreas de maneira individual;
- b) Ajustar pelo método da máxima verossimilhança (MV), modelos com parâmetros compartilhados, separando com distinção da variabilidade do erro entre as unidades amostrais e dentro das unidades amostrais;
- c) Descrever as interpretações biológicas advindas dos resultados dos ajustes;
- d) Selecionar o modelo que melhor se ajusta ao conjunto de dados utilizando critérios não subjetivos que envolvam: melhor predição e parcimônia (AIC), associada a um maior grau de correlação espacial;
- e) Verificar a estrutura da variabilidade e dependência espacial dos volumes nas áreas de estudo, e predizer os volumes por hectare por meio da krigagem simples;
- f) Determinar a qualidade dos ajustes das interpolações obtidas na krigagem pelo método da validação cruzada.

### 3 REFERENCIAL TEÓRICO

#### 3.1 FLORESTA ESTACIONAL DECIDUAL

De acordo com Rambo (1951) encontram-se no estado do Rio Grande do Sul a distinção de três formações características que representam os grandes grupos de florestas, sendo: a Floresta Ombrófila Mista, representada pelas floras tropical e temperada com marcada relevância fisionômica de elementos Coniferales e Laurales; a Floresta Ombrófila Densa, com espécies tropicais afro-brasileiras, muitas evoluídas das áreas serranas próximas ao Oceano Atlântico; e por fim, a Floresta Estacional, distribuídas ao longo da porção meridional do estado, formada principalmente por representantes das floras da Bacia Amazônica e do Brasil.

Segundo Veloso e Góes Filho (1982), o conceito ecológico da Floresta Estacional está ligado ao clima, caracterizado por duas estações distintas: uma chuvosa e outra seca, acentuada variação térmica e uma estacionaridade foliar dos elementos arbóreos dominantes no estrato emergente, os quais têm adaptação à deficiência hídrica ou à queda da temperatura nos meses frios, que atua como seca fisiológica nas espécies tropicais.

Klein (1983) afirmou que a Floresta Estacional pode ser também semidecidual quando possuir uma quantidade de 20 a 50% das folhas do estrato superior trocadas no período de inverno, ao passo que a classificação em Floresta Estacional Decidual, está relacionada à percentagem de folhas trocadas nos meses frios ultrapassando 50%.

É possível diferenciar nesta tipologia três estratos, sendo o primeiro um dossel composto de árvores altas que podem variar de 30 a 35 m de altura e que não formam uma cobertura contínua, ocorrendo como árvores emergentes. O segundo estrato das árvores, comumente entre 20 e 25 m de altura, é bastante denso e contínuo, formado por um número pequeno de espécies características que apresentam valores sociológicos elevados, constituído em sua maioria por espécies perenifólias. O terceiro e último estrato é o das arvoretas, possui um número relativamente elevado de árvores medianas, normalmente entre 5 a 12 m de altura, dentre as quais algumas apresentam valores sociológicos muito elevados (KLEIN, 1983).

### 3.2 ANÁLISE EXPLORATÓRIA DE DADOS (A.E.D)

Segundo Bussab e Morettin (2009), as medidas estatísticas mais utilizadas para expressar a variabilidade de um conjunto de dados são: variância ou desvio padrão; assimetria e curtose; coeficiente de variação e a amplitude total, por serem relativamente fáceis de calcular e interpretar.

Guimarães (1993) mostrou que a variância só poderia ser utilizada de forma isolada, se fossem comparados grupos de dados que possuem unidades e médias iguais. O coeficiente de variação indica o grau de variabilidade do parâmetro em relação aos seus valores médios, não refletindo um significado físico, mas bastante utilizado para comparar quantidades que possuam unidades diferentes. Já a amplitude total leva em consideração apenas os extremos dos valores observados, sendo mais utilizada quando se deseja saber os limites de aplicação de um determinado estimador.

Para Rohlf e Sokal (1969) os valores num ponto são estimados levando-se em conta a média da população e a variabilidade dos valores medidos em torno dessa média. Essa estimativa pressupõe que as amostras são independentes e que possuem distribuição normal. Costa Neto (1977) comentou que é necessário realizar transformações quando os dados originais não atendem aos pressupostos básicos de variância mínima e distribuição normal.

A análise exploratória dos dados é fundamental e obrigatória dentro de qualquer estudo geoestatístico. Essa é uma etapa precursora introduzida para averiguar as propriedades estatísticas e matemáticas dos dados, onde a partir dela, são feitas escolhas do tratamento adequado no qual os dados devem estar sujeitos (FARIAS; LAURENCEL, 2008).

Hoaglin (1992) citou duas técnicas utilizadas na análise exploratória de dados: o dispositivo de ramo e folhas e o gráfico *box plot*. O dispositivo de ramo e folhas possibilita ver quanto o grupo está próximo da simetria e como estão distribuídos os valores, se estão concentrados ou com lacunas em sua distribuição, já o “*box plot*”, apresenta os dados de forma que as observações discrepantes ou *outliers* são facilmente localizadas em sua estrutura. Esse gráfico separa os dados em distâncias denominadas quartis, sendo que o valor do 1º quartil corresponde ao ponto até 25% da distribuição dos dados, o 2º quartil ou mediana, com 50% e o 3º quartil, com 75% dos valores.

### 3.3 VARIÁVEIS REGIONALIZADAS

Tendo como base trabalhos desenvolvidos pelo Engenheiro de Minas Sul-Africano Daniel G. Krige em 1951, Matheron (1971) formalizou a teoria das variáveis regionalizadas, que associou conceitos de dependência espacial e processos aleatórios aos fenômenos observados durante os mapeamentos em campo.

Para Oliveira (2003) esses fenômenos descrevem uma continuidade geográfica que é representada pela tendência que a variável regionalizada tem de que, valores relacionados aos pontos vizinhos são mais diferentes à medida em que são realizadas observações mais distantes.

Fanha (1994) comentou que as variáveis regionalizadas têm características quantitativas, tais como: localização, continuidade, anisotropia e transição, que estão ligadas aos fenômenos naturais que as representam, atuando em um espaço geométrico em que foram definidas e onde será estudada sua variação.

A geoestatística é a ciência que explica a variação de determinado atributo de um lugar a outro no espaço, com certa continuidade. Esses atributos e seus valores estão relacionados com a posição espacial que ocupam, e já que podem tomar valores distintos em diferentes lugares de observação, demonstram certa independência (MATHERON, 1971).

Segundo Landim (1998) a geoestatística é um tópico especial da estatística aplicada que trata de problemas referentes às variáveis regionalizadas, as quais têm um comportamento espacial que mostram características intermediárias entre as variáveis verdadeiramente casuais e as totalmente determinísticas. Elas apresentam uma aparente continuidade no espaço, sendo representadas por funções numéricas ordinárias, que assumem um valor definido a cada ponto no espaço, descrevendo matematicamente um fenômeno natural.

De acordo com Müller *et al.* (1997) valores de um determinado atributo obtidos para uma região de estudo, podem ser classificados e interpolados para localidades em que não foram amostrados, por meio das técnicas de geoestatística, gerando contornos ou zonas de gerenciamento semelhantes representadas por um sistema de mapeamento.

A geoestatística aplicada tem por objetivos detectar na aparente desordem entre as amostras, medidas de relação espacial, realizando estimativas de valores

em locais não amostrados com base em observações conhecidas na população (krigagem), identificando as inter-relações dos atributos no espaço (análises de correlação entre variáveis) e permitindo ainda, estudar padrões de amostras adequadas (VIEIRA, 2000).

### 3.4 HIPÓTESES ESTACIONÁRIAS

Quando se utilizam técnicas geoestatísticas para a análise dos dados, algumas hipóteses de trabalho são assumidas, principalmente as hipóteses estacionárias de 1ª ou de 2ª ordem, assumidas nos dois primeiros momentos de função aleatória (SAMPER; RAMIREZ, 1990).

A hipótese intrínseca ou de 2ª ordem, estabelece que a probabilidade de variação dos valores tomados entre pontos com a mesma distância de separação é igual, ou seja, a relação de dependência espacial é a mesma em qualquer posição de “u” (vetor que separa dois pontos amostrais), assim, presume-se que há uma variação da média, mas admite-se que essa variação é constante dentro dos limites da continuidade espacial (ISAKS; SRIVASTAVA, 1989; VIEIRA; LOMBARDI NETO, 1995).

A estacionaridade de 1ª ordem ou global é a esperança do valor de um ponto no espaço ser igual à média, admitindo-se que a média da amostra é igual à média populacional, ou seja, presume-se que o conjunto dos pontos, para um determinado atributo, tem a mesma média para todos os locais em toda a região de estudo (TRANGMAR *et al.*, 1985).

### 3.5 MÉTODOS DE ESTIMATIVAS EM GEOESTATÍSTICA

Para Olea (1999) a geoestatística é um conjunto de técnicas numéricas para a caracterização de atributos espaciais utilizando basicamente duas ferramentas: modelos probabilísticos, que são usados como forma de lidar com a incerteza associada às variáveis expressas nos estimadores; e técnicas de reconhecimento de padrões associados aos estudos variográficos.



Na geoestatística ocorre uma divisão entre os modelos determinísticos e os modelos estocásticos que estimam variáveis espaciais. Nos determinísticos as distâncias são geométricas e não trazem consigo os erros associados, como por exemplo, o método do inverso do quadrado da distância (IQD), enquanto que os estocásticos são oriundos de processos aleatórios e capazes de quantificar a incerteza associada aos estimadores, no caso, onde se enquadram os modelos geoestatísticos (YAMAMOTO; LANDIM, 2013).

Diggle e Ribeiro Júnior (2007) comentaram que os métodos de estimativa de variáveis correlacionadas no espaço, podem ser separados em: métodos não paramétricos, baseados no ajuste do semivariograma; métodos frequentistas, que trabalham com máxima verossimilhança e, um terceiro baseado em inferência aplicado ao Teorema de Bayes. Apesar dos métodos apresentarem certo grau de relação entre eles, os fluxos dos tratamentos de dados são distintos.

### 3.5.1 Métodos de estimativa por semivariograma

Dentre os métodos utilizados para variáveis correlacionadas no espaço, o semivariograma representa uma das ferramentas mais utilizadas de suporte às técnicas de geoestatística, permitindo representar quantitativamente a variação de um fenômeno regionalizado no espaço, aplicando-se o ajuste aos dados diretamente no semivariograma (HUIJBREGTS, 1975).

Para Huijbregts (1975), o semivariograma experimental é ajustado à curva que proporciona a máxima correlação possível com os pontos plotados. O modelo ajustado é chamado de modelo teórico do semivariograma, e na região onde se observa a continuidade espacial (região estruturada), a geoestatística é aplicada com eficiência.

Conforme Landim (1998) o semivariograma é uma estrutura gráfica que representa medidas de separabilidade dos pares de observações de um determinado fenômeno que se deseja estudar, pois, mede a variância entre os pontos separados por uma distância “u”, de modo que, observações mais próximas possuem uma menor variância que aumenta, à medida que, a distância “u” entre os pares de pontos também aumenta.

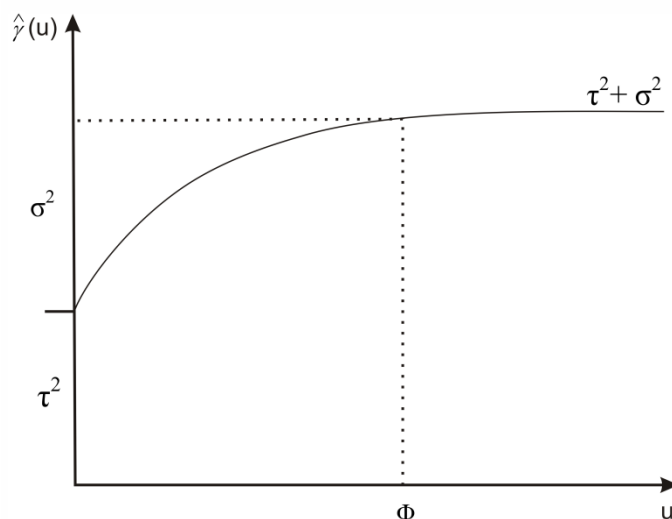


FIGURA 1 – GRÁFICO DO SEMIVARIOGRAMA EXPERIMENTAL COM OS PARÂMETROS EFEITO PEPITA ( $\tau^2$ ), VARIÂNCIA ( $\sigma^2$ ) E ALCANÇE ( $\Phi$ ).  
 FONTE: Adaptado de Diggle e Ribeiro Júnior (2007)

Na Figura 1, o símbolo “ $\hat{\gamma}(u)$ ” representa respectivamente a variância entre os pares de pontos e “ $u$ ”, a distância que separa os pontos. O patamar ou “*still*” ( $\tau^2 + \sigma^2$ ) que representa a soma entre a variabilidade espacial explicada ( $\sigma^2$ ) e o efeito pepita “*nugget*” ( $\tau^2$ ), correspondente a variância máxima atingida pela correlação entre os pares de pontos, assim como, a diferença entre esses dois termos ( $\tau^2 - \sigma^2$ ) obtêm-se a relação sinal-ruído, sendo a variabilidade explicada ( $\sigma^2$ ) o sinal, e o erro aleatório “*nugget*” ( $\tau^2$ ), o ruído.

A dependência espacial ou campo estruturado é dada pelos componentes da variância ( $\sigma^2$ ) e o alcance ou “*range*” ( $\Phi$ ), desse ponto em diante, torna-se invariante a variância entre as amostras, ou seja, cessa a dependência espacial, pois, mesmo com o aumento da distância a função do semivariograma vai oscilar em torno da variância máxima (YAMAMOTO; LANDIM, 2013). O alcance ( $\Phi$ ) mede o grau da correlação espacial entre os pares de pontos, e também o quão forte se dá essa correlação em função do vetor de distâncias (COHEN; SPIES; BRADSHAW, 1990).

O efeito pepita ( $\tau^2$ ) ou “*nugget*” está relacionado com a descontinuidade na origem do semivariograma para distâncias menores do que a menor distância entre as amostras. Podem advir de erros de medição, características da variável ou mesmo, da relação não captada pelo plano amostral ou “*grid*”, sendo impossível quantificar se a maior contribuição provém dos erros de medição ou da variabilidade em pequena escala (ISAACS; SRIVASTAVA, 1989).

Yamamoto e Landim (2013) comentam que a dependência entre estas duas variáveis regionalizadas  $Z(x_a)$  e  $Z(x_a + u)$  representadas pelo seu variograma  $2\hat{\gamma}(u)$ , é uma probabilidade definida como a esperança matemática do quadrado da diferença entre os valores dos pontos relacionados no espaço, separados por um vetor de distâncias “u” também chamado de incremento, representada matematicamente pelo modelo (1):

$$2\hat{\gamma}(u) = E\{[Z(x_a) - Z(x_a + u)]^2\} = \text{var}(x) \quad (1)$$

A divisão por dois da média das diferenças ao quadrado (origem do prefixo “semi”), nos dá a representação do semivariograma (HUIJBREGTS, 1975), podendo ser estimado por meio de uma amostra “ $Z(x_a)$ ,  $a = 1, 2, \dots, n$ ”, expresso conforme o modelo (2):

$$\hat{\gamma}(u) = \frac{1}{2} E\{[Z(x_a) - Z(x_a + u)]^2\} = \frac{1}{2N(u)} \sum_{a=1}^{N(u)} [Z(x_a) - Z(x_a + u)]^2 \quad (2)$$

Onde:

$\hat{\gamma}(u)$  é a variância estimada para cada distância “u”;

$Z$  é a variável aleatória em estudo;

$N(u)$  é o número de pares de valores da expressão “ $Z(x_a) - Z(x_a + u)$ ”, separados por um vetor de distância “u”.

Muitos semivariogramas experimentais podem apresentar o que na literatura geoestatística denomina-se efeito pepita puro, que é a ausência de correlação espacial entre os pares de pontos amostrados. Esse efeito pode ou estar relacionado ao fenômeno espacial em estudo, ou devido à escala de amostragem adotada (YAMAMOTO; LANDIM, 2013) sendo que quando ocorre, a análise por semivariograma não pode ser realizada, sugerindo o uso de outros métodos. O efeito pepita puro não é classificado como um modelo básico, mas aparece como uma constante ( $C_0$ ), na equação do semivariograma e deve ser entendido como  $C_0 = 0$ , quando  $|u| = 0$  (ANDRIOTTI, 1988).

### 3.5.2 Métodos de estimativa por máxima verossimilhança

Basicamente, o que se vê na maioria dos estudos referentes à geoestatística são ajustes gráficos baseados nos métodos de semivariograma, mais utilizados como ferramentas explicativas da variabilidade espacial de um determinado atributo (CARVALHO, 1991; CAMBARDELLA *et al.*, 1994; BERG; KLAMT, 1997; EGUCHI, 2001; ORTIZ, 2002).

Muitas vezes a continuidade espacial não é explícita adequadamente em um semivariograma, levando a resultados imprecisos em relação à distribuição espacial de uma determinada variável em estudo (PLONER; DUTTER, 2000). Na modelagem geoestatística por meio do semivariograma, outro fator importante a ser considerado é que não se assume explicitamente nenhuma distribuição, e os pares de pontos podem se repetir no cálculo das semivariâncias várias vezes, o que se configura em uma desvantagem no método (OLIVEIRA, 2003).

Autores como Diggle e Ribeiro Júnior (2007), Diggle, Tawn e Moyeed (1998) trabalham com a geoestatística baseada em modelos, que descrevem problemas geoestatísticos fundamentados na aplicação de procedimentos estatísticos formais, utilizando como base matemática e estatística, métodos aplicados às funções de máxima verossimilhança, diferentemente dos ajustes gráficos por aproximação aos modelos teóricos de semivariograma.

As vantagens da explicitação de modelos, associada à máxima verossimilhança, são maior eficiência nas estimativas dos parâmetros, e avaliação da incerteza envolvida na estimativa da estrutura de dependência, que podem ser facilmente integradas na previsão, ao contrário da geoestatística baseada em semivariograma (PLONER; DUTTER, 2000).

A probabilidade máxima ou máxima verossimilhança (MV), como é descrita na abordagem frequentista, é uma técnica de grande uso para inferência espacial, incluindo avaliação de significância estatística, cálculo de intervalos de confiança (IC), modelo de avaliação e previsão (MILLAR, 2004).

Jeliazkov e Lee (2010) comentaram que o grande obstáculo da análise de dados multivariados discretos, é o problema em avaliar as probabilidades dos resultados que entram na função de verossimilhança. Os cálculos destas probabilidades envolvem integrais de alta dimensão, fazendo com que os métodos

computacionais associados a essas abordagens na maioria dos casos, tenham que passar por adequações quando implementados para problemas univariados e bivariados.

Diggle e Ribeiro Jr (2007) afirmaram que o método de máxima verossimilhança possui propriedades assintóticas normalmente distribuídas para seus parâmetros, sem viés e com grande eficiência nos estimadores, ao mesmo tempo em que possui flexibilidade para as diferentes transformações da variável resposta, sendo útil para inúmeras aplicações em geoestatística. Segundo Millar (2004) o logaritmo da função da verossimilhança (L-V) é mais utilizado na inferência do que a própria função da verossimilhança, devido a suas qualidades como:

1. Os estimadores de verossimilhança possuem mais eficiência para maximizar a log-verossimilhança do que a função de verossimilhança.
2. Os valores da probabilidade podem ser pequenos na maioria dos casos, mas também podem ser números extremamente grandes, dependendo do modelo testado e do número de dados para estimar a função, gerando cálculos demorados pela máxima verossimilhança em relação ao uso da função log-verossimilhança.

Millar (2004) comentou que a curvatura do logaritmo da verossimilhança, altera em sua inclinação, conforme a derivada de 2ª ordem, tendo fundamental importância tanto na teoria quanto na prática da inferência estatística, pois marca onde o valor é máximo para a função. No contexto da otimização numérica, a função objetivo, é a log-verossimilhança, e a variância estimada é um subconjunto facilmente calculado pelos estimadores da função de máxima verossimilhança.

### 3.6 CRITÉRIO DE INFORMAÇÃO DE AKAIKE (AIC)

Para Barber e Gelfand (2007) a seleção de um modelo estatístico apropriado é de extrema importância, porque a inferência é condicionada no modelo que está sendo especificado corretamente. Harrell e Frank (2002) comentaram que na prática, um modelo estatístico não é uma perfeita descrição dos mecanismos subjacente que geram os dados, mas que seja adequado o suficiente para ser útil, explicando de maneira satisfatória as relações intrínsecas associadas à variável que se deseja responder.

Para Bozdogan (1987), dentre os métodos de seleção para modelos de máxima verossimilhança, o critério de informação de Akaike tem sido amplamente utilizado, devido à sua extensão direta com uma interpretação teórica aplicada ao método de máxima verossimilhança. Segundo esse autor, esta medida baseia-se no conceito de entropia, ou seja, é uma medida relativa das informações perdidas quando um determinado modelo é usado para descrever a realidade.

O AIC foi originalmente desenvolvido por Hirotugu Akaike (AKAIKE, 1974), com o nome de “um critério de informação”, sendo concebido para ser uma medida relacionada à qualidade dos ajustes de um modelo estatístico estimado.

O uso do valor do critério de informação de Akaike (AIC) mostra-se um importante instrumento na escolha de modelos, pois, elimina a subjetividade do processo além de influenciar diretamente na melhoria da confiança dos estimadores associados aos modelos selecionados (VENDRÚSCULO, 2003).

Emiliano (2009) descreveu que o critério de informação de Akaike fundamenta-se na verossimilhança, além de servir para avaliar modelos aninhados (combinados) e também modelos não aninhados, podendo-se comparar quaisquer quantidades de modelos e estudar estruturas de covariâncias associadas aos seus parâmetros.

O valor do AIC para um determinado conjunto de dados não tem qualquer significado prático isoladamente, ele só se torna útil quando é comparado com os valores de AIC de diversos modelos (EMILIANO, 2009). O cálculo do AIC está implementado em alguns pacotes estatísticos tanto comerciais, “S”, “S-Plus”, “SAS”, “SPSS”, “Statistica” entre outros, como em *softwares* de livre distribuição na forma de pacotes, acessados pela interface de programas como o “R” por exemplo.

### 3.7 MODELOS MISTOS E MODELOS MISTOS COMPARTILHADOS

Os modelos mistos ou combinados (aninhados) possuem muitos nomes e aplicações. Podem ser chamados também de modelos hierárquicos, modelos de medidas repetidas, modelos de cluster e ainda, modelos compartilhados. Um modelo misto pode ser utilizado para analisar os dados em series de agrupamentos (*clusters*), ou dados longitudinais, cujo vetor distância está associado ao tempo (DEMIDENKO, 2013).

Para Ramon *et al.* (2006) as metodologias que envolvem a utilização dos modelos mistos vão além da estatística clássica, nessa, é comum supor que as observações retiradas da população em geral são independentes e igualmente distribuídas. Segundo os autores, dados aplicados aos modelos combinados podem ser separados em uma estrutura multinível, possuindo uma hierarquia mais complexa do ponto de vista computacional e matemático.

As observações entre os níveis são independentes, mas dentro de cada “*cluster*” ou agrupamento são dependentes, pois, pertencem à mesma subpopulação, conseqüentemente fala-se de duas fontes de variação, sendo uma entre os *clusters* e a outra dentro dos *clusters*. Uma vantagem dos modelos combinados é a introdução dos efeitos aleatórios ou fixos em vários níveis analisados, ampliando as possibilidades de solução de um problema proposto (RAMON *et al.*, 2006)

O modelo misto é um modelo estatístico não linear, devido principalmente à presença dos parâmetros da variância e, portanto, requer tratamento teórico especial. Em resumo, os modelos mistos podem ser utilizados para as seguintes finalidades (DEMIDENKO, 2013):

- a) Modelar dados com múltiplas fontes de variação;
- b) Modelar dados agrupados ou longitudinais complexos (dados geográficos);
- c) Modelar variabilidade e heterogeneidade biológica;
- d) Como um modelo estatístico para log-verossimilhança;
- e) Fornecer uma base teórica fidedigna para o critério de informação de Akaike (AIC);
- f) Lidar com parâmetros de multidimensionalidade (reconstituição e calibração de imagens geoposicionadas, dentre outras aplicações matemáticas).

### 3.8 PLANO DE AMOSTRAGEM PARA ANÁLISE DA DEPENDÊNCIA ESPACIAL

Quando pretende-se estimar o comportamento do conjunto de todas as observações em potencial referentes a uma população, partindo de um certo número de amostras, é necessário que esses subconjuntos sejam coletados de tal modo que cada observação tenha a mesma chance de ser escolhida. Quando uma amostra é

obtida segundo esse critério denomina-se casual ou aleatório e existem, portanto, diversos esquemas de amostragem, os quais são empregados com o que se conhece a respeito do modelo estatístico, seguindo o objetivo da pesquisa (COCHRAN, 1965).

A realização de uma pesquisa levanta questões relativas ao plano de amostragem de observação da realidade, bem como, a escolha de métodos para processamento e análise de dados. Estudar e planejar essas questões significa buscar opções lógicas e operacionais para a composição de um delineamento que assegurem o valor científico das informações obtidas e das conclusões alcançadas (SOARES, 2000).

De acordo com Campbell e Stanley (1979) entende-se por delineamento, o plano estratégico de observação da realidade que orientará o detalhamento posterior dos métodos e técnicas necessários à execução da pesquisa, sendo categorizadas de acordo com sua natureza lógica e estrutural.

Os delineamentos para alocação de unidades de amostra em estudos de vegetação e inventários florestais basicamente podem seguir três tipos: delineamentos aleatórios, sistemáticos e ainda uma mistura entre ambos. Tais delineamentos possuem vantagens e desvantagens em relação a condicionantes como: instalação, deslocamento, coleta das informações, tipo florestal, precisão, entre outros, porém, o custo do levantamento é um fator comum que relaciona as condicionantes citadas (SCOLFORO; MELLO, 2006).

### 3.8.1 Levantamentos com *grid* regular

Um *grid* regular é um tipo de malha em que as dimensões entre as amostras são de igual tamanho em duas direções perpendiculares (no caso bidimensional), tanto na direção “X” como em “Y” (ANDRIOTTI, 1988).

Yamamoto e Landim (2013) comentaram que a amostragem em *grid* regular (malha de amostras regularmente espaçada), possibilita uma execução em campo mais facilitada, além de não precisar de definições adicionais como distância e direção, contudo deve-se ter em mente que o plano de amostragem reproduza as menores distâncias da relação espacial, representando melhor a variável estudada



em todo experimento. A Figura 2 ilustra uma estrutura de amostras para dados geoestatísticos em um *grid* regularmente espaçado de 40 x 40 m em duas dimensões, onde: “ $X_0$ ” é a posição da variável “ $Z$ ” a ser estimada a partir de  $Z(X_1)$ ,  $Z(X_2)$ ,  $Z(X_3)$  e  $Z(X_4)$ .

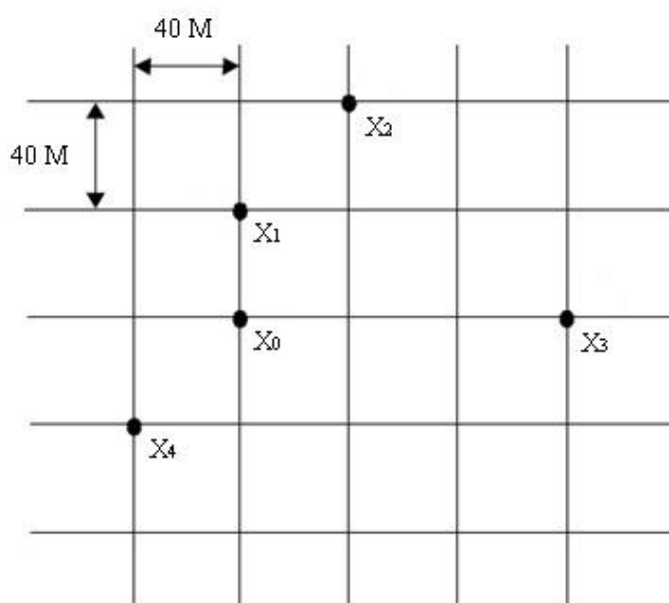


FIGURA 2 – AMOSTRAS COM ESPAÇAMENTO REGULAR DE 40 X 40 METROS, PARA O CÁLCULO DO VALOR DE “ $Z$ ”.

FONTE: O autor (2014)

No exemplo, a amostra na posição  $X_3$  está mais distante que a  $X_2$ , no entanto,  $X_3$  tem mais influência que  $X_2$ , esse fato ocorre, pois,  $X_0$  está sobre a influência direta de  $X_3$ , enquanto  $X_2$  está na verdade mais próxima de  $X_1$ , dessa forma, ao se introduzir as correlações das distâncias na matriz de ponderadores (pesos), evita-se associar pesos erroneamente em amostras menos correlacionadas, diferente de métodos geométricos baseados apenas nas distâncias (ANDRIOTTI, 1988).

### 3.8.2 Levantamentos com *grid* irregular

Chen *et al* (1995) comentaram que o desenho amostral de observações do tipo irregularmente espaçadas geram bons resultados em termos de precisão das

estimativas, porém, em termos de instalação à campo e análise da dependência espacial esse tipo de *grid* é mais complexo.

Em relação a estimativas com pontos irregularmente distribuídos, conforme Yamamoto e Landim (2013) além da distância e da direção, existe uma necessidade de indicar ao modelo parâmetros adicionais, pois, são necessários para a normalização dos pontos obtidos nas amostras. Para cada ponto, tem-se uma janela de interpolação definida pela tolerância angular, passo, direção e largura máxima, onde podem ocorrer um ou mais pontos, ou mesmo nenhum.

Semivariogramas calculados por *grid* irregular geram um maior número de pares de pontos interpolados, resultando em um melhor ajuste ao modelo teórico, assim a curva resultante desse ajuste torna-se mais suavizada (YAMAMOTO; LANDIM, 2013).

De acordo com Deutsch e Journel (1992), considerando o conjunto de amostras irregularmente espaçadas em duas dimensões, neste caso, para determinar o semivariograma experimental é necessário introduzir limites de tolerância para direção e distância. Na Figura 3, a largura de banda (BW) refere-se a um valor de ajuste a partir do qual se restringe o número de pares de observações para o cálculo do semivariograma, a próxima etapa seria o ajuste ao modelo aplicado a uma função teórica.

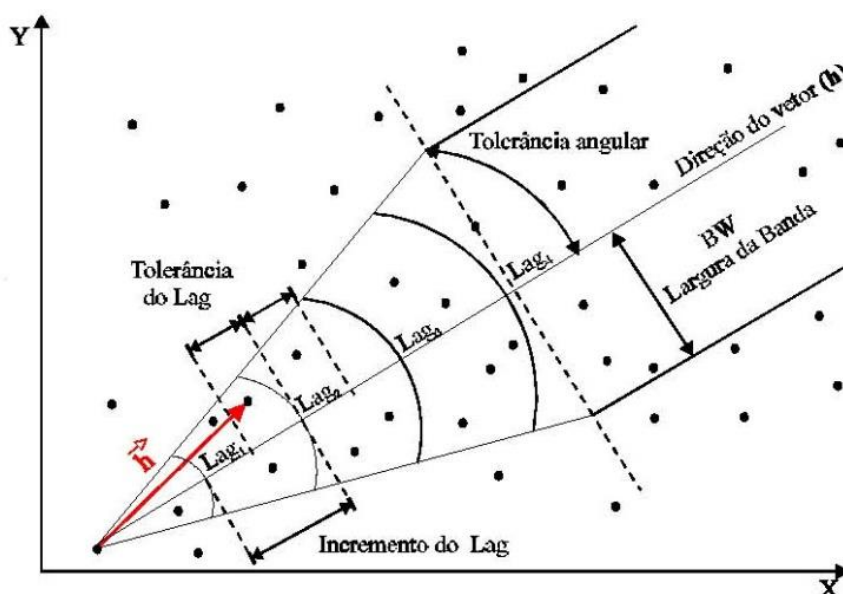


FIGURA 3 - PARÂMETROS PARA O CÁLCULO DO SEMIVARIOGRAMA COM BASE EM AMOSTRAS IRREGULARMENTE ESPAÇADAS EM DUAS DIMENSÕES.  
FONTE: DEUTSCH e JOURNEL (1992)

### 3.9 KRIGAGEM E PREDIÇÃO

A Krigagem é um método de interpolação decorrente da teoria das variáveis regionalizadas, fornecendo estimativas não tendenciosas e com menor variância entre os estimadores não tendenciosos (variância mínima). Esse método foi concebido para expressar médias móveis, que exprimem a variabilidade espacial de um determinado atributo com estimadores não viesados que minimizam os erros de predição (OLIVER; WEBSTER, 1990).

Para Isaaks e Srivastava (1989) a krigagem é o melhor método de interpolação, pois, está baseada na metodologia de melhor estimador linear não tendencioso, o qual é responsável pelo agrupamento dos dados e sua espacialização dentro de uma área pré-definida.

Para determinação dos pesos associados a cada um dos pontos estimadores, o sistema de krigagem baseia-se na ideia de que, quanto maior a covariância entre uma amostra “ $X_i$ ”, em que  $i = 1, 2, 3, \dots, n$ , e o local que está sendo estimado “ $X_0$ ” mais essa amostra deve contribuir para a estimativa nesse ponto (YAMAMOTO; LANDIM, 2013)

Na estimativa pelo método da krigagem o valor desconhecido da variável  $Z_{x_0}^*$  pode ser estimado por uma combinação linear dos “ $n$ ” valores observados, adicionado a um parâmetro  $\lambda_0$ , conforme descritas nas expressões que seguem (JOURNEL, 1988).

$$Z_{x_0}^* = \lambda_0 + \sum_{i=1}^n \lambda_i Z(x_i) \quad (3)$$

Deseja-se um estimador não tendencioso, ou seja, sem viés. Dessa forma, a relação requer que as duas médias sejam iguais para atender a condição, resultando na expressão (4):

$$E [Z_{x_0}] = E [\lambda_0 + \sum_{i=1}^n \lambda_i Z(x_i)] \Rightarrow m = \lambda_0 + \sum_{i=1}^n \lambda_i m \quad (4)$$

Diferente da krigagem simples, a krigagem ordinária não requer o prévio conhecimento da média “ $m$ ”. Neste caso, para que a igualdade da expressão (5) seja satisfeita é necessário que:

$$\lambda_0 = 0 \quad \text{e} \quad \sum_{i=1}^n \lambda_i = 1 \quad (5)$$

Portanto, o estimador de krigagem ordinária é (6):

$$Z_{x_0}^* = \sum_{i=1}^n \lambda_i Z(x_i), \quad \text{com:} \quad \sum_{i=1}^n \lambda_i = 1 \quad (6)$$

Journel (1988) afirmou que minimizando a variância do erro, conforme a expressão:  $(Var[Z_{x_0} - Z_{x_0}^*])$  sob a condição de não viés, os pesos  $\lambda_i$  resultantes são obtidos a partir do cálculo expresso pelo seguinte sistema de equações (7):

$$\sum_{i=1}^n \lambda_i C(x_i, x_j) - a = C(x_j, x_0) \quad \text{para } j = 1, \dots, n \quad (7)$$

Onde:

$C(x_i, x_j)$  e  $C(x_i, x_0)$  são os pares de pontos que se apresentam correlacionados no espaço;

“ $a$ ” corresponde à variância minimizada do erro.

A variância da krigagem ordinária ( $\sigma_{ko}^2$ ) é dada pela seguinte expressão (JOURNEL, 1988):

$$\sigma_{ko}^2 = Var[Z_{x_0} - Z_{x_0}^*] = C_{(0)} - \sum_{i=1}^n \lambda_i C(x_i, x_0) - a \quad (8)$$

O sistema de krigagem ordinária (8) pode ser escrito em notação matricial como (9):

$$K \cdot \lambda = k \Rightarrow \lambda = K^{-1} \cdot k \quad (9)$$

onde, K e k são matrizes das covariâncias e “λ” o vetor dos pesos (10):

$$K = \begin{bmatrix} C_{11} & C_{12} & \dots & C_{1n} & 1 \\ C_{21} & C_{22} & \dots & C_{2n} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ C_{n1} & C_{n2} & \dots & C_{nn} & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix}, \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \alpha \end{bmatrix} \text{ e } k = \begin{bmatrix} C_{10} \\ C_{20} \\ \vdots \\ C_{n0} \\ 1 \end{bmatrix} \quad (10)$$

A variância de krigagem ordinária é dada por (Journel, 1988) (11):

$$\sigma_{ko}^2 = C_{(0)} - \lambda^T \cdot k \quad (11)$$

Yamamoto e Landim (2013) afirmaram que o método da krigagem é utilizado como estimador pontual da variável regionalizada em um determinado local dentro do campo geométrico, com interpolação exata que leva em consideração os pontos de vizinhança para calibrar os estimadores. A krigagem possui variantes tais como: krigagem ordinal; krigagem multigaussiana; lognormal; indicativa e krigagem com equações multiquadráticas, no qual, são escolhidas de acordo com o conjunto de dados, objetivos propostos e a variável que se deseja estudar.

### 3.10 APLICAÇÕES DA GEOESTATÍSTICA NOS LEVANTAMENTOS FLORESTAIS

As variáveis dendrométricas como altura, diâmetro, área basal, biomassa aérea e volume podem apresentar certa dependência espacial, resultante dos processos naturais de crescimento inerentes à competição por espaços e recursos disponíveis em um determinado local. Essa interdependência pode ser analisada através da distância entre as árvores ou fatores relacionados à estabilização das espécies ao longo do desenvolvimento da florestal (WHITMORE, 1990; RODE *et al.*, 2010; AMARAL *et al.*, 2011).

Segundo Mello (2004), levantamentos florestais podem ser incorporados a técnicas geoestatísticas reduzindo o número de amostras para representar a variabilidade de um atributo em determinado local, mantendo assim, a qualidade e a veracidade das informações.

A geoestatística pode contribuir com informações importantes a respeito da dinâmica de crescimento de uma floresta, pois, distúrbios ocorridos impactam a continuidade espacial, alterando a estrutura dos dados, consequência de causas naturais (ventos, deslizamentos, queda de árvores, incêndios, entre outros) ou ação antrópica (corte de espécies de interesse) (MELLO; SCOLFORO, 2000). Para esses autores embora a geoestatística seja mais conhecida na geologia e ciência do solo, suas técnicas possuem grande potencial para serem utilizadas nos inventários florestais, a fim de caracterizar o comportamento espacial de variáveis dendrométricas.

Zawadzki *et al.* (2005) comentaram que durante os últimos anos, aumentou rapidamente a utilização de métodos baseados em informações espaciais com aplicação em inventários florestais, sendo uma tecnologia extremamente importante na determinação de parâmetros biofísicos da floresta e caracterização da estrutura do ecossistema florestal ao nível de parcelas, no entanto, a aplicação destas técnicas para classificação das comunidades florestais complexas não é uma tarefa trivial. A natureza da geoestatística acrescenta um componente estocástico na modelagem de florestas para as aplicações em levantamentos florestais, assim sendo, duas florestas localizadas há uma determinada distância, com a mesma qualidade de sítio e densidade, não têm de ser idênticas, mas provavelmente podem ser, dependendo de sua evolução e dinâmica (NANOS; CALAMA; MONTERO, 2004)

Hernández e Emery (2009) utilizaram abordagem geoestatística para definição de planos de amostragem de baixo custo para avaliação de povoamentos de *Prosopis spp.* na região do deserto do Atacama, norte do Chile, em que era oferecido incentivo fiscal aos proprietários com povoamentos que, após um ano de implantação, tivessem uma taxa de sobrevivência de mudas maior que 75%. Com uso da geoestatística, foi possível dividir a área total em áreas com maior e menor taxa de sobrevivência de mudas, antes realizada por meio de técnicas tradicionais de inventário, distribuindo os recursos oferecidos pelo governo de maneira mais justa e eficiente entre os proprietários.

## 4 MATERIAL E MÉTODOS

Neste estudo, foram utilizadas as três áreas que compõem as fases sucessionais de porte arbóreo trabalhados no projeto “Dinâmica de uma Floresta Estacional Decidual no RS” (LONGHI, 2001), com a coleta dos dados entre os anos de 2003 a 2006.

Os fragmentos localizam-se no município de Santa Tereza, região nordeste do Estado do RS, possuindo cada um aproximadamente três hectares, subdivididos por critérios de homogeneidade fisionômica e com idades atuais determinadas pós-intervenção (relato histórico de remoção seletiva de árvores de interesse econômico) em: Floresta Madura (FMA) com 75 anos; Floresta em Estágio Secundário (FES), com 55 anos e Floresta em Estágio Inicial (FEI) com 30 anos (VACCARO, 1997).

### 4.1 CARACTERIZAÇÃO DA ÁREA DE ESTUDO

Na Figura 4 observa-se o posicionamento das áreas e sua localização geográfica. Os fragmentos FEI e FES estão distantes à aproximadamente 0,8 km entre si, e 3,5 km destes em relação ao fragmento FMA.

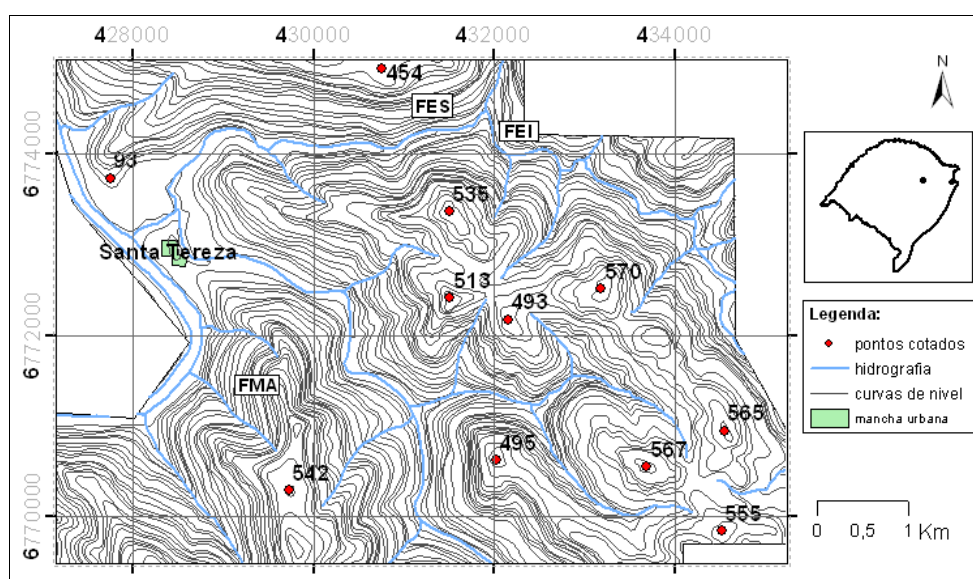


FIGURA 4 – LOCALIZAÇÃO DOS FRAGMENTOS: (FEI) FLORESTA EM ESTÁGIO INICIAL, (FES) FLORESTA EM ESTÁGIO SECUNDÁRIO E (FMA) FLORESTA MADURA  
FONTE: O autor (2014).

#### 4.1.1 Aspectos gerais de clima, solos e relevo.

A Floresta Madura (FMA) possui uma altitude que varia de 255 a 350 m, com 4° de inclinação, exposição sudoeste e relevo ondulado. Já a Floresta Secundária (FES), possui uma altitude que varia de 120 a 350 m, exposição do terreno na direção sudeste e 16° de inclinação média. O relevo foi caracterizado como fortemente ondulado.

A Floresta em Estágio Inicial (FEI) por sua vez possui inclinação média de 5°, exposição do terreno na direção norte, relevo ondulado e altitudes que variam de 125 a 225 m. Na Figura 5 observa-se a exposição do terreno e a variação de altitude nos fragmentos.

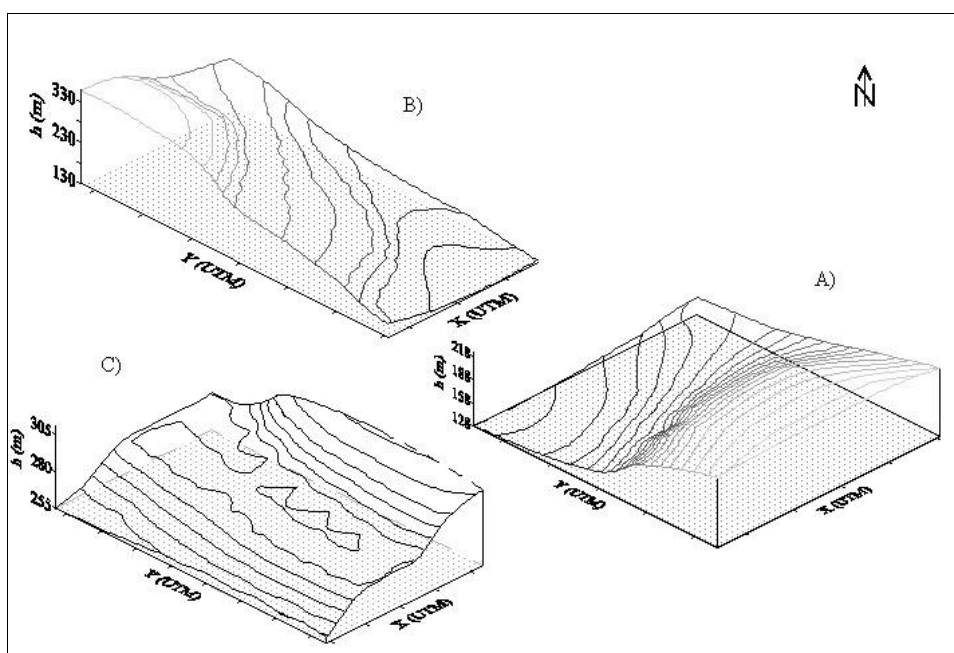


FIGURA 5 - DETALHE DO MICRO RELEVO NAS ÁREAS EXPERIMENTAIS: A) FEI, B) FES E C) FMA.

FONTE: O autor (2014).

O clima da região é classificado como tipo Cfa, subtropical úmido, segundo Köppen (MORENO, 1961), com chuvas bem distribuídas ao longo do ano, temperaturas do mês mais frio entre - 3° e 18° C, e temperatura média do mês mais quente superior a 22° C. Tomando-se as isotermas e as isoetas, a temperatura média anual da região é de aproximadamente 18° C e a precipitação média anual é 1800 mm.



Foram coletadas no centro de cada parcela com trincheiras de 50 cm de profundidade, amostras para análises químicas (espátula) e físicas (anel de Kopecky), num total de 56 coletas realizadas por rotinas e procedimentos metodológicos descritos em Tedesco *et al.* (1995) e EMBRAPA (1997), com a finalidade de caracterizar os tipos de solos existentes nos fragmentos.

A Figura 6 ilustra as quatro principais fases de coleta do solo nas três áreas experimentais.

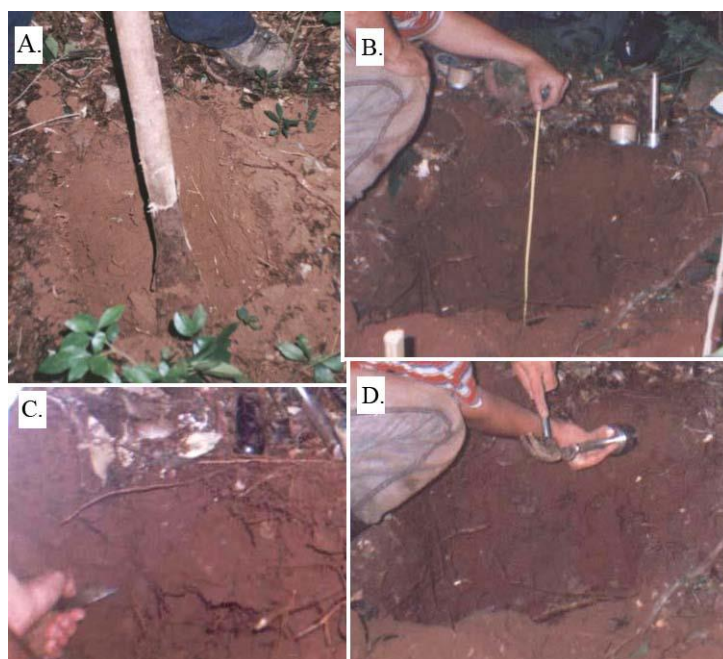


FIGURA 6 - FASES DA COLETA DE SOLO A) ABERTURA DA TRINCHEIRA; B) MEDIÇÃO DA PROFUNDIDADE; C) COLETA DAS ANÁLISES QUÍMICAS E D) COLETA DAS ANÁLISES FÍSICAS.  
FONTE: WOJCIECHOWSKI (2006).

Wojciechowski (2006) realizou a caracterização dos solos predominantes nas mesmas áreas de estudo encontrando as seguintes classes: Chernossolo argilúvico férrico na área de FMA, que apresentaram razoáveis teores de material orgânico, conferindo cores escuras ao horizonte superficial, possuindo também alta fertilidade química (saturação por bases  $\geq 65$  %) e alta capacidade de troca catiônica (CTC) em todo o perfil, e argilúvicos, pois, apresentam um horizonte B textural e férricos pelos elevados teores de ferro ( $\geq 18$  %). Na área FES predominou a classe de solo tipo

Cambissolo Ta eutrófico, diferenciado em função do seu acúmulo de material orgânico no horizonte superficial. O caráter eutrófico refere-se à grande variabilidade nas propriedades químicas, apresentando saturação por bases  $\geq 50\%$ . A classe de solo encontrado na FEI foi o Neossolo Litólico eutrófico, de maior ocorrência no Rio Grande do Sul, apresentam o horizonte “A” ou “O” assentado sobre rocha parcialmente alterada ou inalterada, além de elevada saturação por bases,  $\geq 50\%$  (STRECK; KAMPS; DALMOLIN, 2002).

#### 4.1.2 Aspectos fitofisionômicos

As áreas então localizadas na faixa de vegetação denominada “Fralda da Serra Geral”, com seus limites que vão desde o Rio Ibicuí, no oeste, até o rio Caí, ao leste do Rio Grande do Sul, predominando com mais ou menos 70% a 90% das árvores emergentes perdendo completamente as folhas durante o inverno (REITZ; KLEIN; REIS, 1988).

Para a amostragem da vegetação foram aproveitados no presente estudo os resultados obtidos por Wojciechowski (2006) na mesma área, que encontrou 125 espécies pertencentes a 34 famílias, nos quais 18 exclusivas da FMA, 11 espécies exclusivas da FES e 6 sp. da FEI. A área de FMA apresentou três estratos arbóreos bem definidos e um estrato superior emergente com alguns indivíduos de grande porte, sendo: *Gymnanthes concolor*, *Nectandra megapotamica*, *Sorocea bonplandii* e *Trichilha elegans* as espécies mais frequentes no local.

Notou-se na FES a existência de dois estratos de porte arbóreo contínuo e bem definido, e um terceiro estrato acima em formação, com árvores esparsas e sem entrelaçamento de copas, enquanto que a FEI apresentou dois estratos arbóreos não muito definidos, com um estrato arbustivo denso composto de *Cupania vernalis*, *Trichilha elegans*, *Allophylus edulis* e *Luehea divaricata*, e um herbáceo composto principalmente de gramíneas para ambas as áreas de FEI e FES.

A família Myrtaceae apresentou a maior diversidade de espécies (2 gen., 6 sp.), seguida por Flacourtiaceae (3 gen., 5 sp.), Rutaceae (3 gen., 4 sp.), Lauraceae (2 gen., 4 sp.), Rutaceae (3 gen., 4 sp.) e por último Rubiaceae (4 gen., 4 sp.). Em termos numéricos se destacaram as famílias Sapindaceae, Meliaceae e Fabaceae (WOJCIECHOWSKI, 2006).

## 4.2 OBTENÇÃO DOS DADOS

Foi utilizada uma amostragem do tipo sistemática, distribuindo as unidades amostrais (*grid* de pontos) de forma a cobrir toda a extensão dos fragmentos, mensurando-se um total de 56 unidades amostrais, distribuídas da seguinte maneira: 17 na FEI, 20 na FES e 19 na FMA, alocadas em distâncias regulares de 40 x 40 m, com auxílio de um GPS de navegação e trena. O efeito de borda foi atenuado tomando-se uma distância mínima de 15 metros para o interior dos fragmentos.

Foram adotadas parcelas de área variável com inclusão de árvores proporcional à distância, com intensidade amostral mínima admitida de 10%, calculada a partir da metodologia modificada de Prodan (1965) citado por Péllico Netto e Brena (1997), no qual resultou em unidades com tamanhos entre 200 a 250 m<sup>2</sup>.

No inventário foram anotadas: distâncias, DAP, alturas comercial e total, posições sociológicas, classes de copa, qualidades do tronco, além de identificadas as espécies das oito árvores-amostra com DAP  $\geq 10$  cm mais próximas do centro de cada parcela, sendo a 8ª árvore contada como meia árvore, seguindo um sentido anti-horário de caminamento.

A metodologia original de Prodan para parcelas de área variável com inclusão de seis árvores teve como critério a estabilização do coeficiente de variação da amostragem na parcela, para um menor número de árvores possíveis mantendo a precisão do levantamento. Péllico Netto (comunicação pessoal, 2013) comentou ser possível a utilização dos mesmos princípios de Prodan em outras tipologias florestais, desde que observadas essas mesmas condições.

Nas Figuras 7A, 7B e 7C, observa-se a disposição das parcelas nas áreas de FEI, FES e FMA respectivamente, com círculos de diferentes tamanhos indicando a dimensão da variável volume nas unidades amostrais (U.A), sendo que círculos pequenos indicam menores volumes, aumentando o tamanho dos círculos, proporcionalmente ao aumento do volume por hectare amostrado nas respectivas parcelas. O detalhamento da unidade amostral com uma representação das oito árvores-amostra mais próximas incluídas na parcela de área variável encontra-se na Figura 7D.

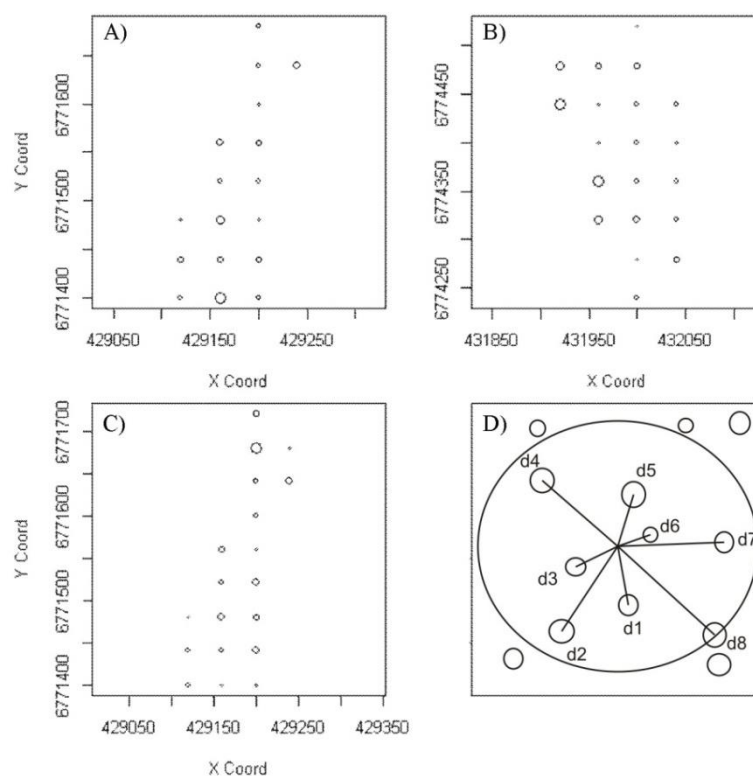


FIGURA 7 - ESQUEMA DE AMOSTRAGEM UTILIZADO NOS FRAGMENTOS: A) FEI; B) FES E C) FMA; D) DETALHE DA UNIDADE AMOSTRAL.  
FONTE: O autor (2014).

O estimador da área basal ( $\hat{G}$ ) de Prodan é similar à ideia de um núcleo ou “Kernel” comumente chamado na literatura, sendo que cada observação é ponderada pela distância referente a um valor central, usando uma relação entre a área basal dentro de um círculo onde o raio é a distância da última árvore de inclusão em relação ao centro da unidade (adaptada para oito árvores), fazendo relação à seguinte proporcionalidade (PELLICO NETO; BRENA, 1997) (12):

$$\hat{G} = \frac{d_1^2 + d_2^2 + d_3^2 + \dots + \frac{d_8^2}{2}}{R_8^2} (2500) \quad (12)$$

Onde:

$\hat{G}$  - estimador de área basal

$R_8^2$  - raio da oitava árvore medida na unidade amostral;

$d_1^2 + d_2^2 + d_3^2 + \dots + \frac{d_8^2}{2}$  - somatório dos diâmetros ao quadrado da primeira, segunda, terceira até a oitava árvore incluída na unidade amostral.

Para o cálculo do número de indivíduos por hectare nas unidades amostrais, têm-se a seguinte expressão (13):

$$N = \frac{(7,5)(10.000)}{\pi R_8^2} = \frac{75.000}{\pi R_8^2} \quad (13)$$

Usando-se o mesmo princípio da proporcionalidade aplicado ao cálculo do número de indivíduos por ha (13), para o volume por ha temos (14):

$$V / ha = \frac{v_1 + v_2 + v_3 + \dots + \frac{v_8}{2}}{\pi R_8^2} (10.000) \quad (14)$$

Onde:

$V / ha$  - volume por hectare;

$v_1 + v_2 + v_3 + \dots + \frac{v_8}{2}$  - somatório dos volumes da primeira, segunda, terceira até a oitava árvore incluída na unidade amostral;

Dessa forma, a estimativa do volume por hectare é dado a partir do somatório dos volumes individuais das árvores-amostra, utilizado um  $f = 0,65$  como fator de forma do volume individual comercial, adotado para florestas decíduais na região serrana do Estado do Rio Grande do Sul (UFSM, 2001).

As análises realizadas do estudo foram construídas utilizando-se os recursos do programa computacional R Core Team (2013) com apoio dos pacotes geoR (RIBEIRO JÚNIOR; DIGGLE, 2001), MASS (VENABLES; RIPLEY, 2002) e SP (PEBESMA; BIVAND, 2005).

### 4.3 ANÁLISE EXPLORATÓRIA DE DADOS

A análise descritiva e exploratória foi realizada a partir da tabulação dos dados e representação gráfica por meio do “*box plot*”, no qual foram descritas as medidas estatísticas usuais como média; mediana, amplitude máxima e mínima dos dados, intervalos interquartílicos e coeficiente de variação total.

Para Vuolo (1996) o *box plot* é um gráfico completo na caracterização de uma variável, que consegue representar cinco medidas estatísticas, em um único conjunto de resultados sendo: os limites interquartílicos, os extremos da distribuição, e a visualização das observações acima dos valores máximos da variável, ou os valores candidatos a “*outliers*”. Foi elaborada a partir da estatística espacial descritiva, uma segunda análise, agrupada em uma janela gráfica de 4 partes sendo:

- A primeira parte descreve os *quartis* subdivididos nos valores observados de acordo com os símbolos: “+” ponto que corta a distribuição do conjunto de dados entre o intervalo que vai do menor valor até 25% do total observado; “Δ” de 25 até 50%; “o” 50 até 75% e “x” de 75% ao valor máximo observado.
- A segunda e terceira parte das análises verificam possíveis tendências na distribuição espacial dos valores de volume observados em relação aos eixos de coordenadas na direção “X (UTM)” e “Y (UTM)”.
- A quarta e última, mostra um histograma de frequência para o volume em m<sup>3</sup> por hectare com objetivo de analisar possíveis assimetrias e também observar como distribuem-se os volumes do conjunto de dados original, e inferir sobre a necessidade e o tipo de transformação aplicada.

Fez-se a transformação dos dados pelo método de *Box-Cox* (BOX; COX, 1964), encontrando como parâmetro de transformação o valor de lambda ( $\lambda$ ) igual à zero. Dessa forma, tem-se a simplificação da variável transformada ( $Y'$ ) do modelo (15), para o modelo (16):

$$Y' = (Y^\lambda - 1) / \lambda, \text{ para } \lambda \neq 0 \text{ e;} \quad (15)$$

$$Y' = \text{LOG} (Y), \text{ para } \lambda = 0 \quad (16)$$

onde:

$Y'$  - variável volume ( $m^3 ha^{-1}$ );

$\lambda$  - parâmetro de transformação lambda.

#### 4.4 DEPENDÊNCIA ESPACIAL DA VARIÁVEL VOLUME ESTIMADA POR VEROSSIMILHANÇA

Foi utilizada a função Matêrn (MATÈRN, 1986) para explicar a estrutura da covariância dos dados de volume nas áreas, bem como, obter os parâmetros do modelo (DIGGLE; RIBEIRO JÚNIOR, 2007) (17).

$$\rho(u) = \{2^{K-1} \Gamma(K)\}^{-1} (u/\Phi)^K K_K(u/\Phi) \quad (17)$$

Onde:

$\rho(u)$  - função de correlação dos pares de pontos observados separados a uma distância “u”;

$\Gamma(K)$  - função Gama;

$u$  - distância entre os pares de pontos;

$\Phi$  - alcance do modelo;

$K_K$  - função Bessel de terceiro tipo de ordem k.

A aproximação da função Matêrn ao modelo de correlação exponencial foi obtida mediante a utilização de um valor do parâmetro de suavização  $kappa = 0,5$ , resultando na expressão (18):

$$\rho(u) = \exp \{-(u/\Phi)^K\} \quad (18)$$

Na realização dos ajustes para os modelos propostos no presente estudo, foram elaborados dois tipos de abordagens para o conjunto de dados. Na primeira abordagem, foram ajustados os modelos às áreas independentes entre si (não compartilhadas) subdivididas em dois tipos: ajuste ao modelo de regressão simples

e ajuste ao modelo não compartilhado, em que suas amostras foram analisadas e modeladas para cada fragmento separadamente.

Uma segunda abordagem descreve os ajustes dos modelos com compartilhamento dos parâmetros através da abordagem *borrow strenght*, subdividida em relação ao parâmetro *nugget* ( $\tau^2$ ) sendo: modelos compartilhados sem *nugget* fixo e modelos compartilhados com *nugget* fixo.

#### 4.4.1 Ajuste de modelos sem componente espacial

Primeiramente foi testada a hipótese em que, a média dos volumes amostrados (modelos ajustados sem a componente espacial) era suficiente para explicar a variabilidade nas áreas em comparação aos modelos de máxima verossimilhança (MV) sem compartilhamento de parâmetros (ajustados para cada área individualmente). Esse ajuste foi realizado com o propósito de verificar se o componente espacial dos modelos de máxima verossimilhança adiciona informação nas estimativas dos parâmetros, ou seja, explicar se existe a dependência espacial para a variável volume nas áreas de estudo.

Para tanto, ajustou-se os modelos sem a componente espacial " $S(X_i)$ " admitindo-se as observações " $y_i$ " do volume na posição " $X_i$ " para as observações da variável volume " $y_i = \mu(X_i)$ ", com a média associada a uma relação linear entre as coordenadas " $X_{1i}$ " e " $X_{2i}$ " da área sendo, " $X_i = (X_{1i}, X_{2i})$ " o vetor da localização espacial dos pares de pontos observados, e " $i = 1, 2, 3, \dots, n$ ", ficando a expressão assim descrita (19):

$$\mu(X_i) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} \quad (19)$$

Onde:

$\mu(X_i)$  – média do processo estocástico;

$\beta_0$  – média geral;

$\beta_1$  e  $\beta_2$  – coeficientes associados as posições no espaço  $X_{1i}$  e  $X_{2i}$  (coordenadas da área).



#### 4.4.2 Ajuste de modelos espaciais não compartilhados

O ajuste de uma função para cada área com os parâmetros, média ( $\beta$ ), *nugget* ( $\tau^2$ ), variância espacial ( $\sigma^2$ ) e alcance ( $\Phi$ ) independentes e não combinados, foi obtido através da incorporação do modelo gaussiano à função da verossimilhança, adicionando uma superfície de tendência linear “ $S(X_i)$ ” que segue a distribuição normal multivariada (NMV) por aproximação (DIGGLE; RIBEIRO JÚNIOR, 2007) (20).

$$S(X_i) \sim NMV[D\beta, \sigma^2 R(\Phi)] \quad (20)$$

Foi utilizado o modelo geoestatístico como base para os ajustes dos volumes em  $m^3 \text{ ha}^{-1}$  por máxima verossimilhança, com as observações  $y_i = y_1, y_2, \dots, y_n$ , normalmente distribuídas, ficando reescrito como (21):

$$y_i = D\beta + S(X_i) + \varepsilon_i, \quad (21)$$

Onde:

$y_i$  - vetor da variável resposta;

$D$  - matriz das variáveis explicativas ou covariáveis;

$\beta$  - vetor dos parâmetros da regressão;

$S(X_i)$  - matriz de efeitos aleatórios dada por  $S(X_i) \sim NMV[0, \sigma^2 R(\Phi)]$ , com média zero e variância  $\sigma^2$ ; e  $R(\Phi)$  é a matriz de covariâncias do elemento  $p(u; \Phi, K)$ .

$\varepsilon_i$  - é um termo independente associado ao erro aleatório com média “0” e variância “ $\tau^2$ ”. Sendo o “ $\tau^2$ ” chamado de “*nugget effect*” ou efeito pepita, “ $I$ ” é a matriz identidade. O termo “ $u$ ” na função de correlação “ $p$ ” é a distância euclidiana entre os pontos observados ( $u = ||x_i - x_j||$ ), em função do grau de associação espacial  $\Phi$ , e  $K$  é o parâmetro de suavização *kappa*.

Os parâmetros da expressão (21) foram estimados pela função do logaritmo da verossimilhança maximizada, a partir do modelo descrito por Diggle e Ribeiro Júnior (2007) (22).

$$L(\beta, \tau^2, \sigma^2, \Phi) = -0,5 \{n \log(2\pi) + \log\{[(\sigma^2 R(\Phi) + \tau^2 I)]\} \\ + (y - D\beta)^T (\sigma^2 R(\Phi) + \tau^2 I)^{-1} (y - D\beta)\} \quad (22)$$

Onde:

$L$  - função log-verossimilhança do modelo geoestatístico;

$\beta, \tau^2, \sigma^2, \Phi$  – parâmetros da média, erro aleatório, variância espacial e alcance da função log-verossimilhança;

$D$  - matriz das covariáveis;

$R(\Phi)$  - matriz das covariâncias do elemento “ $\rho$ ”, em função do alcance ou grau de correlação espacial “ $\Phi$ ”;

$I$  – matriz identidade;

$y$  – variável resposta (volume).

#### 4.4.3 Ajuste de modelos com compartilhamento de força

Uma segunda análise refere-se à ideia de que numa estimativa conjunta, parâmetros de informação do modelo como “ $\beta$ ”, “ $\tau^2$ ”, “ $\sigma^2$ ” e “ $\Phi$ ” de uma determinada área podem “emprestar força” ou informações daquilo em que uma área estiver em comum com a outra. De maneira que a estimativa do modelo possa explicar a variável resposta em outras áreas através de uma distribuição assumida, abordagem conhecida na literatura por “*Borrow Strenght*” (BRILLINGER, 1997; KANGAS; MALTANO, 2006).

Dentro ainda da segunda abordagem, os modelos foram subdivididos em modelos estimados por máxima verossimilhança sem *nugget* fixo e com *nugget* fixo, sendo concebidos para separar a variância entre as unidades amostrais e a variância dentro das unidades amostrais (U.A), ou seja, entre as árvores do mesmo ponto ou *kernel*.

No segundo caso, a variância dentro das unidades amostrais (U.A) é similar à variabilidade não explicada do  $\tau^2$  (erro aleatório ou *nugget*) do modelo geoestatístico. Dessa forma, esse parâmetro teve sua variabilidade determinada *a priori*, através de uma análise de variância (ANOVA), no qual gerou uma estimativa do *nugget* fora do

variograma, reduzindo o número de parâmetros a serem estimados do modelo, de tal modo que, as relações de dependência espacial entre os pontos ocorressem de forma mais nítida do ponto de vista geoestatístico, podendo assim serem estimadas com maior precisão.

Foram utilizadas as mesmas equações (28), (29) e (30), com 1000 iterações e especificação nos parâmetros da função do logaritmo da verossimilhança  $L [\beta, \theta (\tau^2, \sigma^2, \Phi)]$ , em função da combinação das áreas ficando a expressão assim reescrita (23):

$$\begin{aligned} & L\{(\beta^{(i)(j)(k')}), \theta [(\tau^{2(i)(j)(k')}), (\sigma^{2(i)(j)(k')}), (\Phi^{(i)(j)(k')})]\} \\ & = - 0,5\{n^{(i)(j)(k')} \log (2\mu) + \log\{[(\sigma^{2(i)(j)(k')} R(\Phi^{(i)(j)(k')} + \tau^{2(i)(j)(k')} I)]\} \\ & + (y^{(i)(j)(k')} - D\beta^{(i)(j)(k')})^T (\sigma^{2(i)(j)(k')} R(\Phi^{(i)(j)(k')} + \tau^{2(i)(j)(k')} I)^{-1} (y^{(i)(j)(k')} - D\beta^{(i)(j)(k')}) \} \quad (23) \end{aligned}$$

Onde:

$L$  – função log-verossimilhança do modelo geoestatístico;

$^{(i)(j)(k')}$  – Índices que indicam as áreas na ordem FEI, FES e FMA, onde os parâmetros podem ou não ser comuns dependendo das configurações dos ajustes nos modelos testados;

$\beta^{(i)(j)(k')}; \tau^{2(i)(j)(k')}; \sigma^{2(i)(j)(k')}; \Phi^{(i)(j)(k')}$  - parâmetros da média, efeito pepita, variância espacial e alcance para as áreas de Floresta em Estágio Inicial<sup>(i)</sup>, Estágio Secundário<sup>(j)</sup> e Maduro<sup>(k')</sup>, respectivamente.

#### 4.4.4 Construção de modelos compartilhados

Nesse estudo foram selecionados alguns modelos com maior chance de aderência do ponto de vista biológico. Muitos outros modelos foram testados e não entraram no *roll* das análises devido a fatores como explicação biológica pouco provável ou mesmo não convergência dos parâmetros estimados devido a matrizes com covariâncias próximas a zero. Assim, a análise exaustiva de todos prováveis modelos, tal qual é feita na abordagem clássica, ou seja, gerando todas as possibilidades e analisando qual a melhor, frente às estimativas, seria trabalhosa do ponto de vista computacional e analítico.

Em relação ao número de modelos a serem testados, podemos fazer uma analogia a um código binário, onde o mesmo é calculado para uma resolução de 8 bits, sendo  $2^8$  gerando 256 possibilidades para 0 e 1, imaginando a título de suposição que o código tivesse não duas posições, mas sim 4 representando os parâmetros " $\beta$ ,  $\tau^2$ ,  $\sigma^2$  e  $\Phi$ " e os nossos bits fossem a permutação das áreas (111, 112, 121, 122, 123) então, têm-se algo em torno de 2880 modelos candidatos ao ajuste.

Na Tabela 1 observa-se a codificação das áreas para os cruzamentos efetuados nos modelos ajustados, sendo que, na codificação "1,1,1" as áreas são consideradas iguais *a priori* para determinado parâmetro (gera um modelo mínimo), e a codificação "1,2,3", as áreas são consideradas diferentes (gera um modelo máximo). As codificações, "1,1,2"; "1,2,1" e "1,2,2" geram os modelos intermediários.

TABELA 1 - CODIFICAÇÃO DAS POSIÇÕES 111, 112, 121, 122 E 123 EM FUNÇÃO DOS CRUZAMENTOS DAS ÁREAS NA SEQUÊNCIA FEI, FES, FMA

Codificação	Interpretação
1,1,1	FEI = FES = FMA
1,1,2	(FEI = FES) $\neq$ FMA
1,2,1	(FEI = FMA) $\neq$ FES
1,2,2	FEI $\neq$ (FES = FMA)
1,2,3	FEI $\neq$ FES $\neq$ FMA

FONTE: O autor (2014)

Dessa forma, foi proposto o ajuste a um gradiente de modelos baseados na complexidade dos parâmetros " $\beta$ ,  $\tau^2$ ,  $\sigma^2$  e  $\Phi$ " em relação as áreas "FEI", "FES" e "FMA" indo do modelo mais parcimonioso ou modelo mínimo, em que todas as áreas são semelhantes, resultando em um ajuste com menor número de parâmetros estimados, até um modelo mais complexo em que todas as áreas são distintas, como máximo número de parâmetros, de modo que, as combinações advindas dos ajustes em conjunto permitam fazer um melhor uso da informação proveniente das diferentes áreas, ampliando o poder das análises e interpretações biológicas, além de facilitar o entendimento dos aspectos em que as áreas são semelhantes ou diferentes.

#### 4.5 CRITÉRIOS DE SELEÇÃO DOS MODELOS PARA ESTIMATIVA DE VOLUME

Adotou-se como escolha e seleção dos modelos o Critério de Informação de Akaike (AKAIKE, 1974), conhecida pela sigla “AIC”, para as diferentes relações de covariância dos parâmetros combinados, sendo que, o modelo escolhido é aquele que apresenta o menor valor de AIC entre os testados, calculado pela seguinte expressão (24):

$$AIC = - 2*(L-V) + K * Npar \quad (24)$$

Onde:

AIC - Critério de informação de Akaike;

$L-V$  - logaritmo da verossimilhança maximizada;

$K$  - termo da fórmula que penaliza o modelo de maior número de parâmetros em relação a um de menor número ( $k = 2$ ), para a fórmula clássica do AIC;

$Npar$  - número de parâmetros estimados pelo modelo geoestatístico (o número de parâmetros mínimo é 3 e o máximo é 12).

O termo  $(L-V)$  tem uma interpretação oposta ao da função de verossimilhança, e indica que a função com maior verossimilhança terá o menor valor da log-verossimilhança negativa, resultado da mudança de sinal decorrente da transformação dos dados. Esse termo foi utilizado para avaliar o comportamento do máximo da função frente à transformação logarítmica aplicada na variável resposta, no caso, o volume em metros cúbicos por hectare.

Cada modelo foi relacionado com seu respectivo parâmetro ( $\Phi$ ), indicando o maior alcance em metros resultante do compartilhamento de suas matrizes de covariância. Dessa forma, o melhor modelo selecionado deveria necessariamente atender a condição de um menor valor encontrado pelo critério de Akaike, e seu alcance prático ( $AP = 3\Phi$ ) deveria ser superior a menor distância entre os pontos (40 m), em função do uso da equação selecionada na representação da superfície de predição dos volumes por hectare.

Burnham e Anderson (2002) comentaram que o teste de AIC não somente seleciona o modelo mais correto, mas também, compara o quanto é mais correto em relação a diferentes modelos testados, sendo utilizado para qualquer tipo de modelo, tanto lineares, como não lineares, aninhados e não aninhados.

#### 4.6 KRIGAGEM E PREDIÇÃO

Foi realizada a krigagem e posteriormente a confecção das superfícies de predição dos valores estimados de volume para o melhor modelo, ajustado por máxima verossimilhança no método compartilhado e no método de ajuste aplicado às áreas individuais. Adotou-se a krigagem simples (K.S) com a transformada reversa para a escala original, que gerou as superfícies de interpolação, tendo em vista, que o vetor de médias da variável resposta era conhecido (RIBEIRO JÚNIOR; DIGGLE, 2001).

#### 4.7 VALIDAÇÃO CRUZADA

Foi realizada a técnica idealizada por Isaaks e Srivastava (1989) denominada “validação cruzada” para avaliar a capacidade preditiva do modelo com menor valor do AIC e maior alcance prático (AP). A técnica consiste em retirar uma observação do conjunto original dos dados e reajustar o modelo comparando a eficiência da predição em relação ao valor retirado, sendo o processo repetido para todas as observações que compõem as unidades amostrais (U.A). Dessa forma, foram calculadas as medidas relacionadas à dimensão do erro para a variável resposta em  $m^3 \text{ ha}^{-1}$  de acordo com as expressões 25, 26 e 27 descritas (McBRATNEY; WEBSTER, 1986).

a) Erros médios reduzidos (ER):

$$ER = \frac{1}{\eta} \sum_{i=1}^{\eta} \frac{(z_{xi} - \hat{z}_{xi})}{\sigma(\hat{z}_{xi})} \quad (25)$$

b) Desvio-padrão dos erros reduzidos ( $S_{ER}$ ):

$$S_{ER} = \sqrt{\frac{1}{\eta} \sum_{i=1}^{\eta} \frac{|z_{xi} - \hat{z}_{xi}|}{\sigma(\hat{z}_{xi})}} \quad (26)$$

c) Erro absoluto ( $EA$ ):

$$EA = (z_{xi}) - (\hat{z}_{xi}) \quad (27)$$

Onde:

$\eta$  - número de observações;

$z_{xi}$  - valor observado no ponto  $xi$  ;

$\hat{z}_{xi}$  - valor predito pelo estimador geoestatístico da krigagem no ponto  $xi$  sem considerar a observação  $\hat{z}_{xi}$  ;

$\sigma(\hat{z}_{xi})$  - desvio padrão da krigagem no ponto  $xi$  sem considerar a observação  $\hat{z}_{xi}$  .

Para avaliação e interpretação dos resultados de acordo com as expressões 25, 26 e 27, considera-se o melhor ajuste os valores dos erros médios reduzidos ( $ER$ ) e erro absoluto ( $EA$ ) em  $m^3 \text{ ha}^{-1}$  próximos a zero, e um desvio-padrão dos erros reduzidos ( $S_{ER}$ ) próximos a um.

#### 4.8 RESUMO DAS ETAPAS METODOLÓGICAS

As etapas que envolveram a análise da dependência espacial, os tratamentos, processos e ajustes dos modelos propostos, bem como, as estimativas para o volume por hectare, encontram-se ordenados de acordo com o fluxograma em que: MV – máxima verossimilhança; NSM – modelos sem componente espacial;

MV não combinados – modelos de máxima verossimilhança não compartilhados; MV combinados sem  $\tau^2$  fixo – modelos de máxima verossimilhança compartilhados sem *nugget* fixo; MV combinados com  $\tau^2$  fixo – modelos de máxima verossimilhança compartilhados com *nugget* fixo e AED – Análise Exploratória de Dados (FIGURA 8).

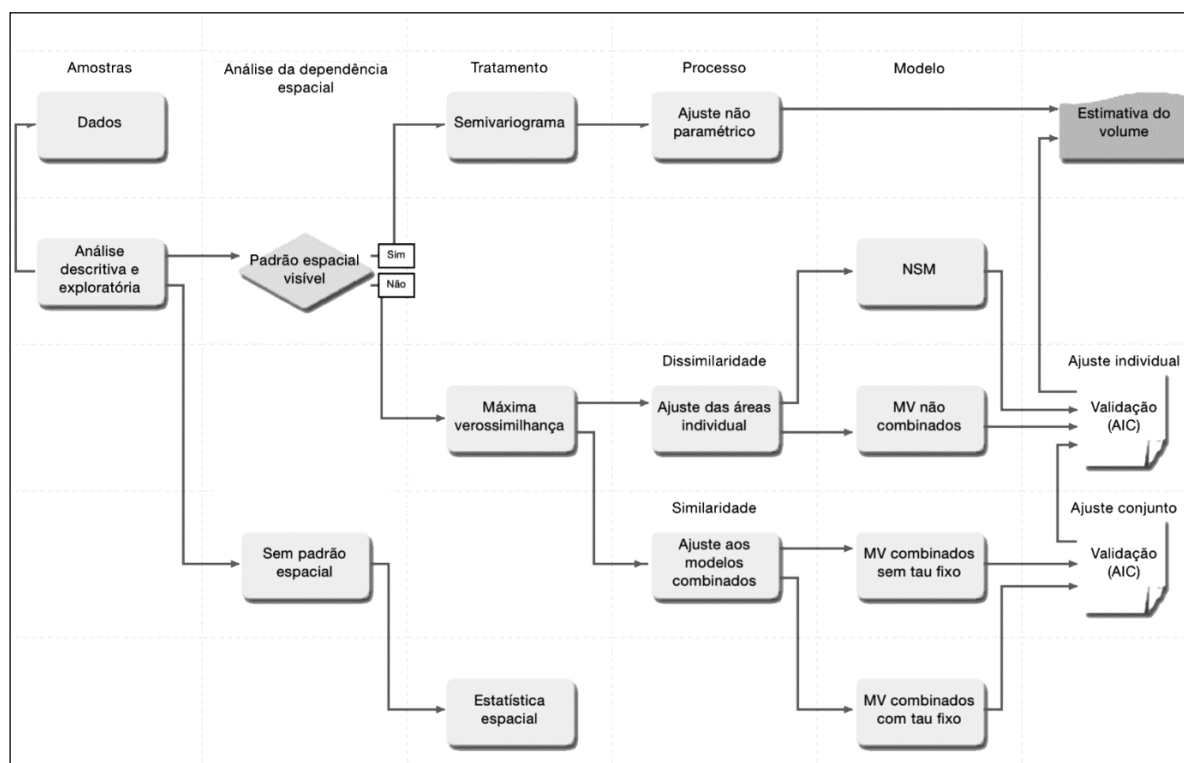


FIGURA 8 – FLUXOGRAMA DAS ETAPAS DA METODOLOGIA DE ANÁLISE E TRATAMENTO DOS DADOS

FONTE: O autor (2014)

Os processos finais da análise que não constam no resumo do fluxograma da Figura 8, incluem a confecção das superfícies de interpolação dos volumes por hectare para cada área, utilizando a krigagem baseada no modelo escolhido e krigagem nas áreas com ajustes individuais, por fim, a validação cruzada com retorno dos erros das predições dos volumes por unidade amostral (U.A) e totais para cada fragmento.

Os valores de volume total em m<sup>3</sup>/ha do inventário encontrados pelo método geoestatístico (Krigagem) foram comparados aos valores obtidos pelo estimador de Prodan nas unidades amostrais (U.A), com a finalidade de verificar o desempenho geral das estimativas de superfície.



## 5 RESULTADO E DISCUSSÃO

A seguir, são apresentados os resultados obtidos para estatística descritiva e exploratória com distinção para os dados originais e transformados, o ajuste e a seleção dos modelos não compartilhados, sendo: modelos sem componente espacial (NSM) e modelo não compartilhado (MV), assim como, os modelos compartilhados ajustados por máxima verossimilhança com e sem *nugget* fixo. Por fim, a krigagem oriunda do melhor modelo construído pelas estimativas dos parâmetros em conjunto nas três áreas experimentais, e a krigagem para as áreas ajustadas em separado.

### 5.1 ESTATÍSTICA DESCRITIVA E ANÁLISE EXPLORATÓRIA

Nas Figuras 9A, 9B e 9C observa-se o comportamento da distribuição e simetria do conjunto de dados em relação aos valores médios do volume por hectare. Pode-se notar nos gráficos de *box plot* que ocorreram valores localmente atípicos nas três áreas de estudo, resultando em assimetria do conjunto de dados em relação aos valores médios na FEI, FES e FMA respectivamente, indicativo de que não houve normalidade das distribuições. Mello (2004) comenta que esses valores atípicos influenciam negativamente nos ajustes obtidos por métodos de semivariogramas, devendo mediante análise, serem removidos quando possível.

Na área de Floresta em Estágio Inicial (FEI), os valores mínimos e máximos da distribuição foram respectivamente  $26,3 \text{ m}^3$  e  $125,0 \text{ m}^3$ , com média  $56,8 \text{ m}^3$  por hectare, divididos em intervalos interquartílicos de 40,2; 50,2 e  $66,6 \text{ m}^3$  (FIGURA 9A) respectivamente para 25%, 50% e 75% do total observado.

Na Floresta em Estágio Secundário (FES), os valores mínimos e máximos da distribuição foram  $25,1 \text{ m}^3$  e  $249,0 \text{ m}^3$  respectivamente, com média de  $90,3 \text{ m}^3$  por hectare, com intervalos interquartílicos de 42,4; 73,1 e  $111 \text{ m}^3$  por hectare, respectivamente para 25%, 50% e 75% dos dados (FIGURA 9B).

A Floresta em Estágio Avançado (FMA) teve valores observados mínimos e máximos para o volume por hectare de  $10,8 \text{ m}^3$  e  $265,0 \text{ m}^3$  respectivamente, com

média igual a 82,0 m<sup>3</sup> e distribuição dos dados de 37,8; 66,1 e 115,1 m<sup>3</sup>, respectivamente para as distâncias do 1º, 2º e 3º quartil (FIGURA 9C). As áreas tiveram valores de coeficiente de variação (C.V) máximo de 30% para os volumes observados por hectare.

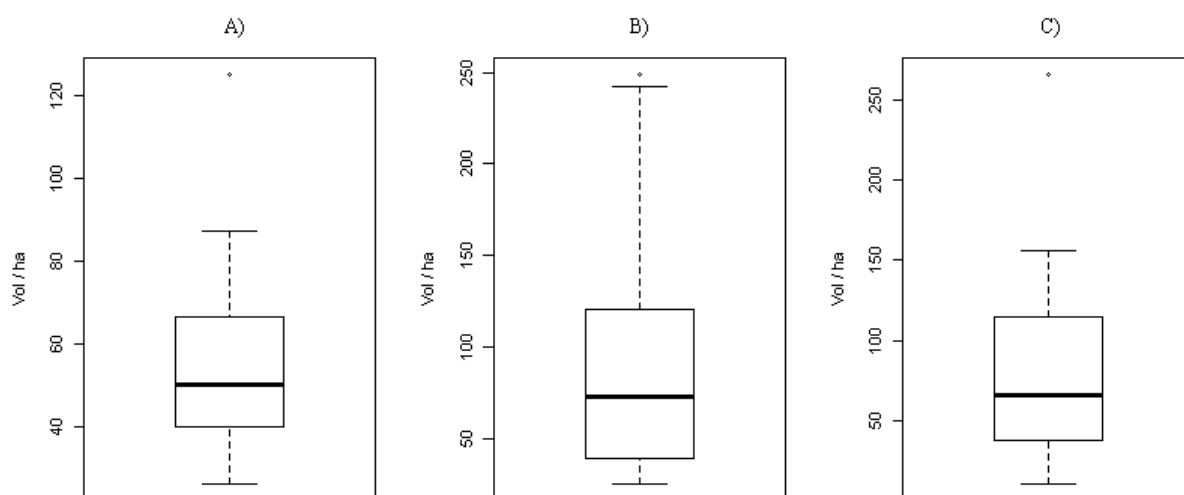


FIGURA 9 - GRÁFICOS *BOX PLOT* DA DISPERSÃO DOS VOLUMES (m<sup>3</sup> ha<sup>-1</sup>) NAS ÁREAS DE A) FEI; B) FES E C) FMA.

FONTE: O autor (2014)

A Figura 10 apresenta um conjunto de gráficos da análise exploratória para os dados da Floresta em Estágio Inicial (FEI), onde inicialmente (FIGURA 10A), observam-se regiões semelhantes das parcelas representadas pelos símbolos (quartis de uma distribuição empírica dos valores medidos do volume por hectare), com tendências mais nítidas nas Figuras 11A e 12A.

Nos gráficos de tendência dos valores dispostos no espaço para as direções das coordenadas “x” e “y”, Figuras 10B e 10C respectivamente, nota-se que o comportamento da variável volume na área FEI ocorreu de maneira homogênea, sendo essa, uma condição importante para a aplicação da hipótese intrínseca nesse local.

Em análise do histograma (FIGURA 10D), podemos observar que as maiores densidades para os volumes por hectare, foram encontrados entre as classes 40,0 e 60,0 (m<sup>3</sup> ha<sup>-1</sup>), caracterizando uma curva de distribuição com assimetria positiva para os valores observados.

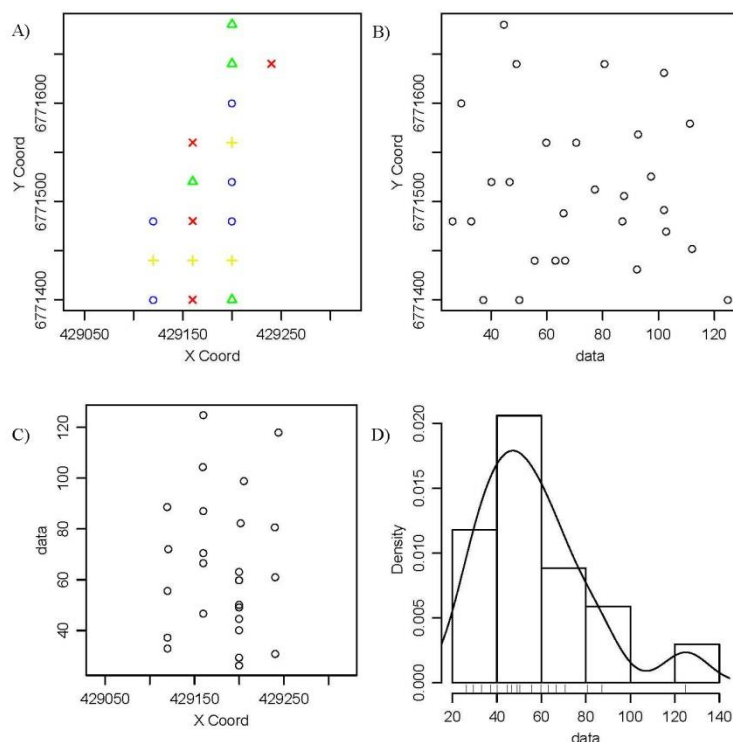


FIGURA 10 – A) PARCELAS SUBDIVIDIDAS EM *QUANTIS*, DISPERSÃO DOS VALORES CONTRA AS COORDENADAS B) X E C) Y, D) HISTOGRAMA (VOL/HA) NA ÁREA DE FEI.  
FONTE: O autor (2014)

Na análise preliminar do fragmento FES (FIGURA 11A) em relação à dispersão dos dados contra o eixo de coordenadas “x” (FIGURA 11B) e “y” (FIGURA 11C), observa-se que ocorreu uma distribuição homogênea nos valores dos volumes na área, sem tendências evidentes em nenhum sentido. Em relação ao histograma dos dados na FES, percebe-se que ocorreu uma maior concentração dos volumes entre o intervalo de 0 a 100  $\text{m}^3 \cdot \text{ha}^{-1}$  (FIGURA 11D). Vieira (1995) a respeito da distribuição não uniforme das observações, comenta que tal fato pode gerar semivariogramas sem patamar, ou com patamar não definido, os chamados variogramas transitivos, dificultando a aplicação dos ajustes por semivariograma.

Na Floresta em Estágio Avançado (FMA) também foi observada uma distribuição regular dos volumes em relação aos eixos “x” (FIGURA 12B) e “y” (FIGURA 12C), com adensamento da variável resposta entre os intervalos de 0 a 50 ( $\text{m}^3 \cdot \text{ha}^{-1}$ ) e de 100 a 150 ( $\text{m}^3 \cdot \text{ha}^{-1}$ ) (FIGURA 12D), apresentando falhas na distribuição oriunda de classes de volumes faltantes, padrões típicos de florestas que sofreram cortes seletivos de alguns indivíduos de maior interesse comercial (SCHNEIDER; FINGER, 2000).

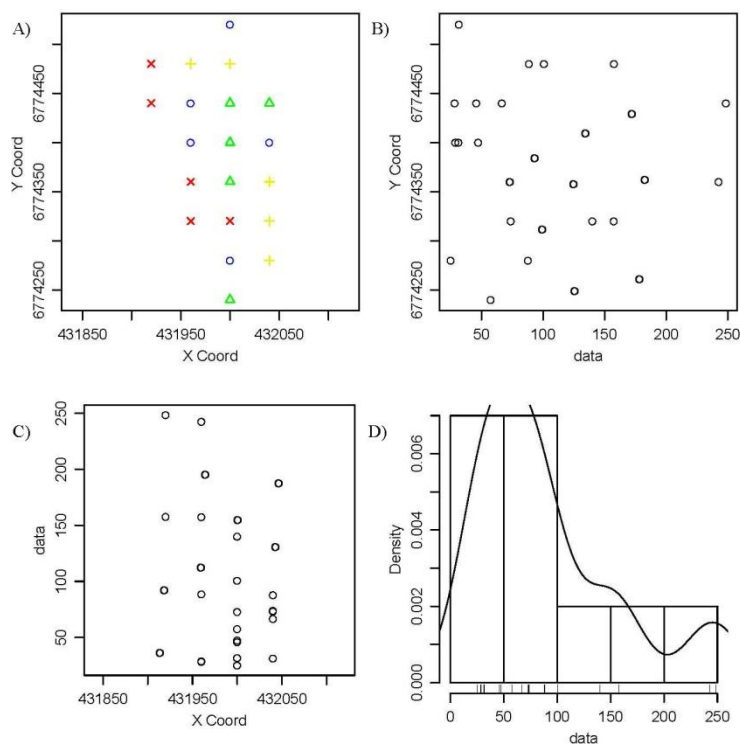


FIGURA 11 – A) PARCELAS SUBDIVIDIDAS EM *QUANTIS*, DISPERSÃO DOS VALORES CONTRA AS COORDENADAS B) X E C) Y, D) HISTOGRAMA (VOL/HA) NA ÁREA DE FES.  
 FONTE: O autor (2014)

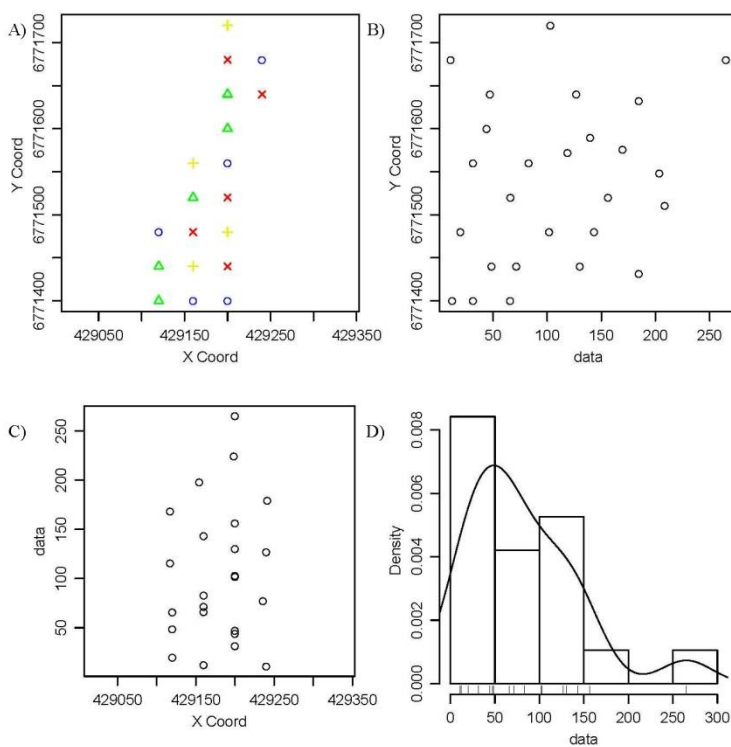


FIGURA 12 – A) PARCELAS SUBDIVIDIDAS EM *QUANTIS*, DISPERSÃO DOS VALORES CONTRA AS COORDENADAS B) X E C) Y, D) HISTOGRAMA (VOL/HA) NA ÁREA DE FMA.  
 FONTE: O autor (2014)

A análise dos histogramas das Figuras 11D, 12D e 13D indicaram que a distribuição dos dados não é normal e sua transformação se faz necessária, tendo em vista a aplicabilidade dos métodos de ajustes por máxima verossimilhança empregados no estudo. Dessa forma, foi realizada uma análise da maximização do logaritmo da verossimilhança (*L-V*) através do gráfico de *Box-Cox* (BOX; COX, 1964), onde foi encontrado um valor de  $\lambda$  comum para as três áreas ( $\lambda = 0$ ), indicando a necessidade da transformação dos dados. Schneider e Finger (2000) comentaram que geralmente a transformação dos dados originais acarreta em redução da variabilidade e normalização dos resíduos.

Ribeiro Júnior (1995) comentou que a normalidade dos dados é uma exigência na aplicação das técnicas de modelagem com métodos que utilizam ajustes por máxima verossimilhança, no qual tem por base a distribuição normal Gaussiana.

Um valor de  $\lambda$  comum obtido na análise gráfica pela maximização da função do logaritmo da verossimilhança foi importante, pois, facilitou a abordagem conjunta nas áreas possibilitando dessa forma, a construção e a aplicação dos modelos de máxima verossimilhança com parâmetros compartilhados. Os resultados da transformação *Box-Cox* podem ser visualizados nos gráficos dos perfis da verossimilhança (FIGURA 13).

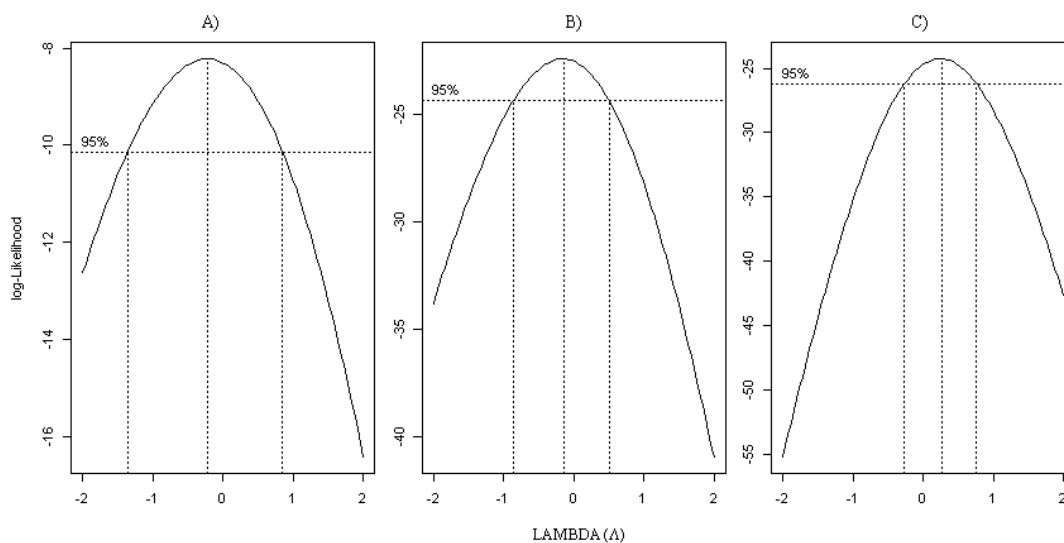


FIGURA 13 - PERFIL DA LOG-VEROSSIMILHANÇA PARA O PARÂMETRO DE TRANSFORMAÇÃO BOX-COX APLICADO A VARIÁVEL VOLUME NAS ÁREAS: A) FEI, B) FES E C) FMA.  
FONTE: O autor (2014)

Para Vuolo (1996) a transformação obtida em geral melhora aproximação dos valores observados em relação a distribuição normal, porém, é necessária cautela, pois, não há garantias de que mesmo a melhor escolha da transformação a ser realizada, produzirá um conjunto de dados transformados que seja adequado à suposição de normalidade.

Kitanidis (1997) comentou ser possível que alguns dos valores originalmente fora do conjunto de dados, acompanhem a distribuição depois de aplicada uma transformação. Na prática é encontrar uma boa maneira de descrever a distribuição assimétrica dos dados, utilizando transformações que resultem em uma distribuição mais simétrica.

Em relação aos resultados obtidos por uma transformação, Kitanidis (1997) ressaltou ainda que esses devem ser cuidadosamente examinados, pois em muitos casos, os chamados valores extremos são as medidas mais interessantes no conjunto de dados que descrevem um fenômeno, essa recomendação de fato vale para qualquer transformação usada em modelos.

A transformação dos dados reduziu a variância da distribuição, que no caso dos dados transformados, acompanha a média. Os resultados dos ajustes podem ser observados nos gráficos de *quantil-quantil* (q-q) da Figura 14, que demonstram os quantis empíricos dos resíduos versus os quantis teóricos das observações dos volumes por hectare nas áreas de estudo.

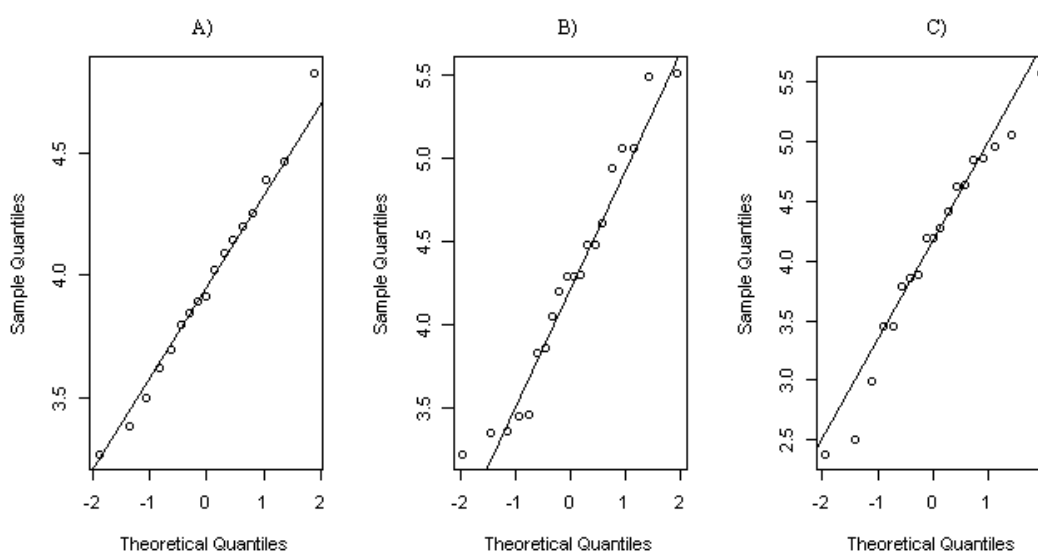


FIGURA 14 - GRÁFICOS QUANTIL-QUANTIL (Q-Q) DA DISTRIBUIÇÃO NORMAL PARA O VOLUME NAS ÁREAS EXPERIMENTAIS: A) FEI, B) FES E C) FMA.  
FONTE: O autor (2014)

## 5.2 AJUSTES DE MODELOS GEOESTATÍSTICOS POR MÁXIMA VEROSSIMILHANÇA PARA ESTIMATIVA DO VOLUME.

Nos itens que seguem, são apresentados os resultados numéricos dos ajustes aplicados nos modelos geoestatísticos pelo método da máxima verossimilhança logarítmica, de acordo com a metodologia do estudo e especificados nas Tabelas 2 a 15, sendo ordenados estruturalmente conforme o fluxograma descrito na metodologia do estudo (item 4.8, FIGURA 8).

Os modelos não espaciais (NSM) não constam nas tabelas dos modelos ajustados individualmente, pois, foram incorporados nos resultados para melhor comparação e discussão no item 4.3.1, da tabela de seleção do AIC para os modelos não combinados para estimativa de volume.

### 5.2.1 Ajuste de modelos não compartilhados

Os ajustes dos parâmetros dos modelos não compartilhados mostram que nas três áreas, a Floresta em Estágio Secundário (FES) foi à única que teve sua dependência espacial explicada, tendo seu alcance prático ultrapassando a menor distância entre os pontos (44,03 m), porém, nesse caso, essa dependência foi explicada até muito próxima do limite do *grid* de pontos (40,00 m).

TABELA 2 - ESTIMATIVA DE MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME PARA OS MODELOS DE PARÂMETROS NÃO COMPARTILHADOS NAS ÁREAS FEI-FES-FMA

Áreas	Estimativas					AP
	$\beta$	$\tau^2$	$\sigma^2$	$\tau^2 + \sigma^2$	$\Phi$	
FEI	3,995	0,151	0,005	0,156	2,4747	7,43
FES	4,273	0,001	0,474	0,475	14,697	44,03
FMA	4,101	0,685	0,025	0,710	2,546	7,62

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático ( $3\Phi$  para *Kappa* 0,5).

De acordo com os resultados da Tabela 2 pode-se afirmar que em uma estimativa individual, apenas um dos modelos ajustados foi possível gerar informação suficiente para explicar a variabilidade dos volumes, no caso, o modelo ajustado para a Floresta em Estágio Secundário (FES). Enquanto que as áreas FEI e FMA tiveram sua dependência espacial ( $3\Phi$ ) de 7,43 e 7,62 m, respectivamente. Esse resultado é um indicador de que as unidades amostrais tratadas em separado (sem compartilhamento de parâmetros), não possuem informação suficiente para explicar a dependência espacial nos locais não amostrados, e a solução poderia estar na combinação das U.A de cada área em uma distribuição maior, possivelmente oriunda dos modelos compartilhados.

Nesse aspecto, em uma abordagem geoestatística que tem a preocupação de investigar a variável, não somente do ponto de vista descritivo, mas também a sua distribuição no espaço, a informação em relação aos volumes das outras áreas não seria válida para os processos de interpolação e predição, prejudicando a análise.

## 5.2.2 Ajuste de modelos compartilhados

Nesse item, são descritos os resultados dos ajustes pelo método de máxima verossimilhança aplicado aos dados transformados, considerando os efeitos nos modelos sem *nugget* fixo (variabilidade entre as U.A) e com *nugget* fixo (variabilidade dentro das U.A), pela abordagem dos modelos compartilhados.

### 5.2.2.1 Ajuste de modelos compartilhados sem *nugget* fixo

Na Tabela 3 observa-se o ajuste do modelo mais parcimonioso ou modelo mínimo em relação ao número de parâmetros. Este modelo foi testado partindo do pressuposto de que todas as áreas são semelhantes, ou seja, possuem valores muito próximos de seus volumes médios ( $\beta$ ), variância não espacial ( $\tau^2$ ) bem como os componentes pertencentes à variabilidade espacial explicada ( $\sigma^2$ ) e o alcance ( $\Phi$ ). Em uma análise preliminar, partindo da premissa básica que os fragmentos pertencem à mesma fitofisionomia, diferindo somente em idade, a resposta do grau



de associação espacial entre os pares de pontos das unidades amostrais do modelo, ajustou um alcance prático (AP = 9,15 m) abaixo do *grid* de 40 x 40 m, com uma variabilidade explicada dos volumes até essa distância. Outro ponto a ser observado é a relação sinal-ruído, que é a razão da variância espacial ( $\sigma^2$ ) pela variância aleatória ( $\tau^2$ ), considerada fraca nesse modelo que não é apropriado para predições devido aos motivos citados (BONAT; RIBEIRO JUNIOR; ZEVIANI, 2011).

TABELA 3 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ ,  $\tau^2$ ,  $\sigma^2$  e  $\Phi$ ) DO MODELO NAS ÁREAS FEI-FES-FMA (MODELO MÍNIMO).

Modelo 01	Estimativas					AP
	$\beta$	$\tau^2$	$\sigma^2$	$\tau^2 + \sigma^2$	$\Phi$	
FEI						
FES	4,1168	0,0070	0,4939	0,5009	3,0493	9,15
FMA						

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático (3 $\Phi$  para *Kappa* 0,5).

Na Tabela 4 observa-se o ajuste de um modelo que estimou 8 parâmetros, com a possibilidade das áreas combinarem a mesma informação referente aos seus componentes de estrutura e variabilidade espacial ( $\sigma^2$  e  $\Phi$ ). Da mesma forma o modelo testado admite que cada área tenha seus próprios valores de média ( $\beta$ ) e erro aleatório ou nugget ( $\tau^2$ ).

TABELA 4 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\sigma^2$  e  $\Phi$ ) DO MODELO NAS ÁREAS FEI-FES-FMA

Modelo 02	Estimativas					AP
	$\beta$	$\tau^2$	$\sigma^2$	$\tau^2 + \sigma^2$	$\Phi$	
FEI	3,9599	0,0020		0,3226		
FES	4,2648	0,0020	0,3206	0,3226	5,2162	15,65
FMA	4,1014	0,3838		0,7044		

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático (3 $\Phi$  para *Kappa* 0,5).

Analisando os resultados das estimativas dos parâmetros para esse cenário, houve uma melhora do parâmetro no alcance do modelo “01” de ( $AP = 9,15$  m), para 15,65 do modelo “02” ( $AP = 15,65$  m), contudo, ainda não se configura em um modelo adequado, pois, apresenta um alcance ( $\Phi = 5,21$ ) insuficiente para cobrir a informação dos volumes nos fragmentos.

Observa-se que essas duas tentativas “02” e “01”, resultaram em uma perda do grau de dependência espacial encontrado no modelo ajustado por máxima verossimilhança não compartilhada para a área de Floresta em Estágio Secundário (FES), sendo que, o alcance do modelo reduziu de 44,03 para 15,65 e 9,15, respectivamente nos modelos “02” e “01”. Conforme observado, as áreas podem tanto compartilhar força dos seus parâmetros, como também, reduzir o alcance (grau de dependência espacial), de acordo com a configuração do modelo testado.

Wu (2010) comenta que, em se tratando de modelos mistos, a combinação resultante da tentativa de ajuste dos parâmetros, modelando os efeitos fixos e aleatórios, tem resultado direto na matriz de variância e covariância dos modelos.

Isto é possível, pois os modelos foram construídos de forma a garantir que as áreas analisadas expressem suas similaridades e também suas dissimilaridades, dando a mesma chance para ambas as probabilidades ocorrerem. A análise final resultante fica em torno dos critérios de escolha em função do objetivo do estudo.

Os modelos que seguem, ditos intermediários somente a título de separação, abordam essas combinações, não só dos parâmetros em relação às áreas, mas também o compartilhamento entre as áreas (FEI-FES, FEI-FMA e FES-FMA), no intuito de que áreas mais “fortes” em termos de amostragem, possam contribuir com mais informações nas áreas com baixa amostragem, no qual se configura um dos objetivos principais do método compartilhado.

Observa-se na Tabela 5 o modelo intermediário “03” com oito parâmetros, compartilhando as médias ( $\beta$ ), o nugget ( $\tau^2$ ), a variância espacial ( $\sigma^2$ ) e o alcance ( $\Phi$ ) entre os fragmentos FEI e FES, com separação desses parâmetros para área de Floresta Madura (FMA). Esse ajuste tenta relacionar duas das áreas que são mais próximas, no caso a áreas FEI e FES (0,8 km) em relação à área FMA, mais distante das demais (3,5 km). A proximidade das áreas segregadas em todos os parâmetros no modelo testado resultou em uma correlação espacial, expressa no alcance prático na ordem de 39,52 m para FEI e FES, e 10,31 m para FMA.

Esse resultado demonstra que mesmo áreas mais próximas, podem apresentar alguma característica biológica que as diferencie (expressa na forma dos parâmetros analisados) resultando em uma fraca dependência espacial, no qual, os modelos com parâmetros combinados podem de forma isolada ou conjunta analisar.

TABELA 5 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ ,  $\tau^2$ ,  $\sigma^2$  e  $\Phi$ ) DO MODELO NAS ÁREAS FEI-FES

Modelo 03	Estimativas					AP
	$\beta$	$\tau^2$	$\sigma^2$	$\tau^2 + \sigma^2$	$\Phi$	
FEI						
FES	4,1247	0,0069	0,3439	0,3508	13,1749	39,52
FMA	4,1014	0,2039	0,6892	0,8931	3,4366	10,31

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático (3 $\Phi$  para *Kappa* 0,5).

Observa-se em relação ao ajuste obtido pelo modelo “03”, que houve um aumento da correlação espacial entre os pares de pontos amostrados expresso nos valores do alcance prático (AP) em relação aos modelos anteriores (15,65 para 39,52 m), porém, abaixo da menor distância entre o *grid* de pontos (40 x 40 m).

Dessa forma tem-se que os fragmentos FEI-FES, em conjunto, nessa configuração do modelo, não explicam a variabilidade dos volumes em toda a extensão das áreas, da mesma forma que, a modelagem conjunta pode unir e “emprestar força” de uma área para outra, também podem, devido à integração de suas covariâncias, diminuir ou “enfraquecer” a explicação espacial de uma área “mais favorecida” em termos de amostragem para outra “menos favorecida”.

Ao truncar os parâmetros do modelo para determinada hipótese, inclui-se uma condição ou inferência *a priori*, para analisar seus efeitos no modelo *a posteriori*. O resultado do processo é analisado de forma integrada, porém, antes dos testes de validação, um parâmetro importante a ser observado durante a discussão dos modelos ajustados, é o alcance prático, pois ele representa uma distância real em campo, e seu resultado é dado pela multiplicação do valor do alcance ( $\Phi$ ) por 3, quando *Kappa* = 0,5 (modelo exponencial).

Diggle e Ribeiro Júnior (2007) comentaram que o alcance ( $\Phi$ ) e o parâmetro de suavidade  $Kappa$  ( $k$ ) do modelo Matérn (EXPRESSÃO 17) não são ortogonais, significa que, as escalas correspondentes às diferentes ordens de correlação Matérn, não são diretamente comparáveis, mas uma depende da outra. Para tal, suas aproximações são dadas para o alcance prático em  $3\Phi$ ,  $4,75\Phi$  e  $5,9\Phi$ , parametrizadas com  $kappa$  igual a 0,5; 1,5 e 2,5 respectivamente, e  $\sqrt{3}\Phi$  para função de correlação gaussiana.

Foi testado um modelo intermediário com nove parâmetros ajustados “04” (TABELA 6), compartilhando as médias ( $\beta$ ) e o nugget ( $\tau^2$ ) para as áreas FEI e FES, e fixando o alcance ( $\Phi$ ) do modelo nas áreas de FES e FMA. Como interpretação biológica tem-se que, as áreas FEI e FES podem ser semelhantes em seus volumes médios, com uma variabilidade espacial ( $\sigma^2$ ) diferente para as 3 áreas e incertezas associadas ao erro ( $\tau^2$ ) igual nas áreas FEI e FES. Essa combinação ajustou-se bem aos dados, com uma distância em que as amostras se encontram correlacionadas de 66,09 m nas áreas FEI e FES. Yamamoto e Landim (2013) comentaram que o alcance garante que todos os vizinhos situados dentro de um círculo com esse raio (alcance) são tão semelhantes que podem ser usados para estimar valores para qualquer ponto entre eles.

TABELA 6 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$  e  $\tau^2$ ) DO MODELO NAS ÁREAS FEI-FES E ( $\Phi$ ) NAS ÁREAS FES-FMA

Modelo 04	Estimativas					AP
	$\beta$	$\tau^2$	$\sigma^2$	$\tau^2 + \sigma^2$	$\Phi$	
FEI	4,0155	0,0007	0,1585	0,1592	0,5043	1,51
FES			0,5374	0,5381	22,0324	66,09
FMA	4,1014	0,7097	0,1211	0,7098		66,09

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático ( $3\Phi$  para  $Kappa$  0,5).

Na Tabela 7 nota-se um modelo com nove parâmetros sendo as médias ( $\beta$ ) comuns para as áreas de FEI e FES, a variabilidade associada ao erro aleatório ( $\tau^2$ ) diferente para cada área, mas, com a variabilidade espacial estruturada ( $\sigma^2$  e  $\Phi$ ) igual para FES e FMA.

O modelo “05” resultou em um alcance ( $\Phi$ ) menor que a menor distância entre os pontos (11,78) e um alcance prático igual a 35,37 m para as áreas FES e FMA. Na Floresta em Estágio Inicial (FEI) o alcance prático ficou bem abaixo da relação do *grid* de pontos (7,62 m).

TABELA 7 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ ) DO MODELO NAS ÁREAS FEI-FES E ( $\sigma^2$  e  $\Phi$ ) NAS ÁREAS FES-FMA

Modelo 05	Estimativas					AP
	$\beta$	$\tau^2$	$\sigma^2$	$\tau^2 + \sigma^2$	$\Phi$	
FEI	4,0394	0,0195	0,1426	0,1620	2,5388	7,62
FES		0,0070	0,5084	0,5154	11,7897	35,37
FMA	4,0980	0,2169		0,7253		35,37

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático (3 $\Phi$  para *Kappa* 0,5).

Ajustou-se o modelo “06” com nove parâmetros levando em conta o compartilhamento de “ $\beta$ ,  $\tau^2$  e  $\Phi$ ” para as áreas FEI-FES, e o parâmetro da variabilidade espacial ( $\sigma^2$ ) diferente para as três áreas. Essa condição ajustou um modelo que é forte candidato à predição em uma primeira análise, visto que o alcance prático (AP) foi o maior dentre os modelos testados anteriormente (AP = 75,54 m). Sendo confirmado o modelo “06” pelos critérios de seleção propostos, a espacialização dos volumes se daria apenas em duas áreas, ficando a FMA sem informação em uma possível predição nessa área, pois o alcance prático (0,57 m) é muito pequeno a ponto de influenciar localmente a média. Percebe-se até agora a ausência de um modelo compartilhado que explique a variabilidade espacial acima do limite mínimo do *grid* de pontos (40 m) em todas as áreas (TABELA 8).

Pélissier e Goreaud (2001) consideram que a formação de padrões em sistemas biológicos é normalmente mais complexa, pois, possuem como característica forte a heterogeneidade, visto que, métodos clássicos de estudos de padrões e distribuição do ponto de vista espacial costumam assumir hipóteses de homogeneidade e independência das observações, o que de fato nem sempre ocorre.

TABELA 8 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ ,  $\tau^2$  e  $\Phi$ ) DO MODELO NAS ÁREAS DE FEI-FES

Modelo 06	Estimativas					AP
	$\beta$	$\tau^2$	$\sigma^2$	$\tau^2 + \sigma^2$	$\Phi$	
FEI	4,0154	0,1582	0,0700	0,2282	25,1811	75,54
FES			0,4723	0,6305		
FMA	4,1014	0,0597	0,7055	0,7652	0,1932	0,57

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático (3 $\Phi$  para  $Kappa$  0,5).

O modelo “07” com dez parâmetros, supõe que as três áreas possuem em comum a variância espacial ( $\sigma^2$ ) dos volumes de seus indivíduos distribuídos no espaço, diferindo em suas médias ( $\beta$ ), erro aleatório ( $\tau^2$ ) e alcance ( $\Phi$ ). Conforme observar-se na Tabela 9, a condição testada resultou em um alcance prático de 3,43; 46,69 e 9,21 m, respectivamente para FEI, FES e FMA.

TABELA 9 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO O PARÂMETRO ( $\sigma^2$ ) DO MODELO NAS ÁREAS FEI-FES-FMA

Modelo 07	Estimativas					AP
	$\beta$	$\tau^2$	$\sigma^2$	$\tau^2 + \sigma^2$	$\Phi$	
FEI	3,9599	0,0009	0,1557	0,1562	1,1442	3,43
FES	4,2682	0,3164		0,04926	15,5642	46,69
FMA	4,1014	0,5539		0,7153	3,0692	9,21

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático (3 $\Phi$  para  $Kappa$  0,5).

O modelo “08” com 12 parâmetros testou um ajuste *a priori* em que todas as três áreas possuem características distintas, ou seja, sem estipular efeitos fixos, considerados somente os efeitos aleatórios e deixando os parâmetros livres para as estimativas das médias dos valores observados dos volumes por hectare, variabilidade espacial, erro aleatório e grau de dependência espacial, distintos para todas as áreas.

Em análise desse ajuste, as áreas “mais fracas”, ou seja, com menor número de amostras (FEI-FMA) perderam parte do alcance para FES, tendo seu alcance aumentado (AP = 44,10 m). Na Tabela 10 observa-se que o modelo teve alcances práticos de 0,83; 44,10 e 0,63 m para FEI, FES e FMA respectivamente, quando comparado com os modelos não compartilhados (TABELA 2), os resultados indicam que os efeitos aleatórios (todas as áreas são diferentes) resultantes do ajuste nos modelos combinados podem ser muito próximos aos obtidos para as áreas separadamente (modelagem individual). Sendo que nos modelos não compartilhados, o fragmento de maior número de amostras “FES” teve um alcance prático de 44,03 m, e as áreas FEI e FMA 7,43 e 7,62 m respectivamente.

TABELA 10 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COM PARÂMETROS ( $\beta$ ,  $\tau^2$ ,  $\sigma^2$  e  $\Phi$ ) LIVRES NAS ÁREAS FEI-FES-FMA (MODELO MÁXIMO)

Modelo 08	Estimativas					AP
	$\beta$	$\tau^2$	$\sigma^2$	$\tau^2 + \sigma^2$	$\Phi$	
FEI	3,9600	0,0005	0,1556	0,1561	0,2763	0,83
FES	4,2733	0,0012	0,4722	0,4734	14,7003	44,10
FMA	4,1014	0,0042	0,7054	0,7096	0,2104	0,63

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático ( $3\Phi$  para  $Kappa$  0,5).

### 5.2.2.2 Ajuste de modelos compartilhados com *nugget* fixo

O erro aleatório intra-parcela, ou seja, dentro das unidades amostrais (U.A), foi modelado pelo logaritmo da verossimilhança fora do semivariograma, para posterior adição desse efeito do erro fixo nos modelos, através de uma análise de variância (ANOVA), usando como fator o logaritmo dos volumes, e o quadrado médio dos resíduos como estimativa do erro ( $\tau^2$ ). Dessa forma, os resultados dos valores fixos do *nugget* obtidos fora do semivariograma foram 1,597; 2,001 e 3,076, respectivamente para a Floresta em Estágio Inicial (FEI), Floresta em Estágio Secundário (FES) e Floresta Madura (FMA).

No primeiro modelo utilizando o *nugget* calculado fora do semivariograma “09” foram compartilhadas as médias ( $\beta$ ), variabilidade espacial ( $\sigma^2$ ) e alcance  $\Phi$ , considerados dessa forma, semelhante para as três áreas, resultando em um ajuste com três parâmetros (modelo mínimo).

Na análise do processo de ajuste do modelo “09”, observa-se que a correlação dos parâmetros para essa hipótese resultou em um baixo grau de dependência espacial nas áreas, evidenciado pelo baixo valor obtido no alcance prático (AP = 1,21 m). Este resultado indica que tanto as médias, bem como, a variabilidade espacial dos volumes ou até mesmo o próprio alcance ( $\Phi$ ), mostram-se diferentes para um dado parâmetro ou mesmo área (TABELA 11).

TABELA 11 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$ ,  $\sigma^2$  e  $\Phi$ ) DO MODELO NAS ÁREAS FEI-FES-FMA (MODELO MÍNIMO-*NUGGET* FIXO)

Modelo 09	Estimativas			AP
	$\beta$	$\sigma^2$	$\Phi$	
FEI				
FES	4,1125	0,2111	0,4028	1,21
FMA				

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático ( $3\Phi$  para *Kappa* 0,5).

Calegario *et al.* (2009) estudando o uso de modelos não-lineares de efeito misto em dois níveis com efeitos fixos e aleatórios (erro nas unidades amostrais e nas árvores dentro das unidades) para estimativas de alturas com dados desbalanceados, constataram que a inclusão de tais níveis reduziu significativamente o erro associado aos estimadores do modelo.

Na Tabela 12, foram compartilhados os parâmetros ( $\beta$  e  $\sigma^2$ ) para FEI, FES e FMA, e o alcance ( $\Phi$ ) nas áreas FEI e FES com o *nugget* fixo, resultando em um modelo com quatro parâmetros estimados. O grau de dependência espacial para esta combinação resultou em boas estimativas dos parâmetros de média e variância espacial, porém com AP de 44,11 m na FEI e FES, e 8,12 m no fragmento FMA.



TABELA 12 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$  e  $\sigma^2$ ) DO MODELO NAS ÁREAS FEI-FES-FMA E ( $\Phi$ ) NAS ÁREAS FEI-FES (NUGGET FIXO)

Modelo 10	Estimativas			AP
	$\beta$	$\sigma^2$	$\Phi$	
FEI			14,7026	44,11
FES	4,1126	0,2118		
FMA			2,7082	8,12

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático ( $3\Phi$  para  $Kappa$  0,5).

Observa-se na Tabela 13 o ajuste das áreas FEI-FES que compartilham das mesmas médias e alcance ( $\beta$  e  $\Phi$ ), com uma associação da variabilidade espacial ( $\sigma^2$ ) em relação ao volume nas áreas FES-FMA, resultando em uma equação ajustada com 6 parâmetros. Na interpretação biológica do modelo “11”, observa-se que as áreas FEI e FES combinaram seus parâmetros explicativos, obtendo um alcance prático de 82,50 m ( $\Phi$  = 27,49) nas áreas FEI e FES, com uma variância ( $\sigma^2$ ) ou a distribuição espacial dos volumes comum entre FES-FMA.

Gregoire e Schabenberger (1994) utilizando uma técnica similar ao *nugget* fixo para efeitos mistos de modelos não lineares por máxima verossimilhança restrita, e também por meio de equações de estimativas generalizadas (GEE) para predizer o volume de árvores individuais, encontraram melhorias nas estimativas dos modelos quando separadas a variância residual aleatória da variância total.

TABELA 13 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO OS PARÂMETROS ( $\beta$  e  $\Phi$ ) DO MODELO NAS ÁREAS FEI-FES E ( $\sigma^2$ ) NAS ÁREAS FES-FMA (NUGGET FIXO)

Modelo 11	Estimativas			AP
	$\beta$	$\sigma^2$	$\Phi$	
FEI		0,1296	27,4921	82,50
FES	4,0121			
FMA	4,1015	0,3675	3,4926	10,47

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático ( $3\Phi$  para  $Kappa$  0,5).

A Tabela 14 demonstra o resultado dos estimadores para o modelo “12” ajustado com oito parâmetros, testando a “combinação de força” da variabilidade espacial ( $\sigma^2$ ) entre FES-FMA, e do parâmetro alcance ( $\Phi$ ) entre FEI-FMA. O resultado dessa cadeia de covariâncias, mostra que apenas a área FES teve um alcance prático satisfatório (AP = 47,47 m).

A condição de que FEI e FMA pudessem compartilhar da mesma característica de estrutura espacial, gerou um baixo alcance resultante ( $3\Phi = 6,14$  m), resultado que evidencia as diferenças em relação à ocupação dos espaços, entre uma floresta em fase inicial em relação a uma em estágio avançado de desenvolvimento. Kent e Dress (1979) comentaram que a dinâmica de mortalidade e ingresso em florestas jovens e maduras, podem influenciar a formação de padrões espaciais devido a ação competitiva de seus vizinhos no decorrer do desenvolvimento de uma população.

TABELA 14 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COMPARTILHANDO O PARÂMETROS ( $\sigma^2$ ) DO MODELO NAS ÁREAS FES-FMA E ( $\Phi$ ) NAS ÁREAS FEI-FMA (NUGGET FIXO)

Modelo 12	Estimativas			AP
	$\beta$	$\sigma^2$	$\Phi$	
FEI	3,9598	0,008	2,0476	6,14
FES	4,2714	0,3253	15,7309	47,47
FMA	4,1013		2,0476	6,14

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático ( $3\Phi$  para  $Kappa$  0,5).

Observa-se na Tabela 15 o resultado do ajuste considerando as áreas diferentes para suas médias ( $\beta$ ) e variância estruturada ( $\sigma^2$  e  $\Phi$ ) na condição do *nugget* fixo ( $\tau^2$ ), resultando no modelo “13” com 9 parâmetros estimados (modelo máximo).

Semelhante às situações anteriores em que o modelo foi testado com seus parâmetros livres, ou seja, sem fixar os efeitos *à priori* no modelo, a área FES obteve maior alcance prático (AP = 45,91 m), tendo os fragmentos FEI e FMA atingido alcances práticos de 1,14 e 5,29 m, respectivamente.

TABELA 15 - ESTIMATIVA DA MÁXIMA VEROSSIMILHANÇA APLICADA À VARIÁVEL VOLUME COM OS PARÂMETROS ( $\beta$ ,  $\sigma^2$  e  $\Phi$ ) DO MODELO NAS ÁREAS FEI-FES-FMA (MODELO MÁXIMO-NUGGET FIXO)

Modelo 13	Estimativas			AP
	$\beta$	$\sigma^2$	$\Phi$	
FEI	3,9599	0,0001	0,3818	1,14
FES	4,2705	0,2746	15,3065	45,91
FMA	4,1014	0,4022	1,1659	5,29

FONTE: O autor (2014).

NOTA:  $\beta$  = log da média (vol.ha<sup>-1</sup>);  $\tau^2$  = efeito pepita;  $\sigma^2$  = variância espacial;  $\tau^2 + \sigma^2$  = patamar;  $\Phi$  = alcance; AP = alcance prático ( $3\Phi$  para  $Kappa$  0,5).

### 5.3 SELEÇÃO DO MODELO PELO CRITÉRIO DE AKAIKE (AIC).

Foram separados os modelos não compartilhados (sem componente espacial e com componente espacial) e compartilhados (sem *nugget* fixo e com *nugget* fixo), para uma melhor explanação e visualização das diferenças entre os métodos testados.

#### 5.3.1 Modelos não compartilhados para estimativa do volume por hectare

Na Tabela 16 observa-se o resultado do teste de AIC para os modelos sem o componente espacial (NSM) e os modelos não combinados ajustados individualmente para cada área. Observa-se que todos os modelos lineares sem componente espacial tiveram um valor do teste do AIC menor que os modelos ajustados individualmente para cada área. Mesmo com os acréscimos de dois parâmetros a mais aos modelos, os valores do teste foram próximos, ficando a diferença entre os métodos, nesse caso, apenas relacionada ao número de parâmetros.

Tanto os modelos lineares sem componente espacial, quanto os modelos ajustados por máxima verossimilhança, tiveram uma relação de continuidade espacial para os volumes em metros cúbicos por hectare abaixo da distância mínima entre os pontos das U.A, com exceção da área FES a qual teve um alcance prático de 44,03 m.

TABELA 16 - SELEÇÃO DOS MODELOS COM AJUSTE LINEAR SEM COMPONENTE ESPACIAL (NSM) E POR MÁXIMA VEROSSIMILHANÇA (MV) NAS ÁREAS FEI-FES-FMA PELO CRITÉRIO DE AKAIKE (AIC)

Modelos	Áreas	L-V	Npar	AP	AIC
NSM	FEI	-75,66	2	8,30	155,3
	FES	-111,9	2	44,03	227,8
	FMA	-105,1	2	0,01	214,3
MV	FEI	-75,66	4	8,30	159,3
	FES	-111,9	4	44,03	231,3
	FMA	-105,1	4	0,01	218,3

FONTE: O autor (2014).

NOTA: L-V = log-verossimilhança; Npar = número de parâmetros do modelo, AP = alcance prático ( $3\Phi$  para  $Kappa$  0,5); AIC = Critério de informação de Akaike.

### 5.3.2 Modelos compartilhados para estimativa do volume por hectare

A Tabela 17 mostra o ranking dos modelos ajustados por máxima verossimilhança logarítmica em ordem crescente de AIC. O modelo que melhor se ajustou ao conjunto dos dados entre todos os métodos testados foi o modelo “11” com um AIC de 120,10. O grau de correlação espacial foi o dobro da menor distância entre os pontos ( $\Phi = 27,4921$ ), que teve um alcance prático de 82,50 m nas áreas FEI e FES, ficando a FMA com um alcance menor ( $AP = 10$  m), resultantes dos efeitos fixos e aleatórios aplicados à matriz de covariância desse modelo.

Os valores do AIC para o modelo “12” foram muito próximos do melhor modelo, no entanto, esse modelo possui mais parâmetros que o modelo “11”. Outro ponto que deve ser levado em consideração quando se trata de modelos combinados, é que o modelo deve explicar a dependência espacial da variável em estudo no maior número de locais possíveis. Tratando-se de áreas compostas, o alcance ( $\Phi$ ) é um indicativo importante como resultado do processo de escolha, sendo também um dos critérios adotados para predição da superfície utilizando a interpolação por krigagem.

Observando os modelos com melhor AIC, no modelo “12”, a correlação entre as amostras foi maior na FES com um valor do alcance prático de 47 m, enquanto que nas áreas FEI e FMA atingiram 6,14 m.

Comparando em relação ao primeiro modelo, nota-se que a dependência espacial na condição que foi testada, explicou a variabilidade dos volumes em apenas uma das áreas (FES), enquanto que o modelo de melhor AIC “11” explicou a dependência em duas áreas (FEI-FES).

Para os modelos combinados sem *nugget* fixo, o melhor valor de AIC encontrado foi no modelo “01”, porém o grau de dependência espacial dado pelo alcance nas áreas foi pequeno ( $AP = 9$  m), indicando que a correlação entre os pontos na condição dessa igualdade para as áreas é fraca. Os modelos “04” e “06” aparecem como boas alternativas para relação de dependência espacial, com valores de alcance prático de 66 m (FES-FMA) e 75m (FEI-FES) respectivamente, e ainda valores de AIC próximos ao modelo “01”.

A análise em conjunto do grau de dependência espacial somado ao AIC é importante, pois, facilita a escolha do modelo em relação à interpolação da superfície por krigagem e a representação da distribuição dos volumes por hectare nas áreas, um dos objetivos do estudo.

Ainda referente à Tabela 17, nota-se que não houve dentre os modelos compartilhados um grau de dependência espacial para os volumes em metros cúbicos por hectare nas três áreas acima do *grid* mínimo da amostragem, o que indica que a variabilidade espacial possa estar ocorrendo em distâncias menores que o *grid*, em pelos menos uma ou até duas das três áreas dependendo do modelo.

Comparando os valores de AIC, entre os métodos de ajuste não compartilhados e compartilhados para estimativa de volume, nota-se que a combinação das áreas obteve sucesso em melhorar as estimativas dos parâmetros para todos os modelos. Pois o maior valor do AIC dos modelos compartilhados (08,  $AIC = 129,77$ ), é inferior ao menor valor encontrado nos modelos não compartilhados (Nsm,  $AIC = 160,2$  m).

Referente ao resultado, Gregoire e Schabenberger (1994) afirmaram que a utilização de covariáveis como altura, ou mesmo variáveis de distância, podem melhorar substancialmente a interpretação dos efeitos mistos nos modelos ajustados por métodos de máxima verossimilhança, melhorando os valores de precisão do AIC.

Da mesma forma, comparando os resultados para os métodos compartilhados de ajuste, os modelos com *nugget* fixo foram superiores em relação aos sem *nugget*

fixo. Isso evidencia que houve um ganho em relação às estimativas dos parâmetros e consequentemente nos valores de AIC, quando modelado o efeito aleatório dentro da unidade amostral.

O modelo “08” ajustado com doze parâmetros, que relaciona a média dos volumes, o *nugget*, a variabilidade espacial e o alcance como sendo características individuais de cada área, resultou no valor mais alto no critério do AIC (129,77). Esse modelo de acordo com os critérios estabelecidos de parcimônia, maximização da função log-verossimilhança e grau de dependência espacial teve uma relação fraca do ponto de vista da distribuição dos volumes nos locais amostrados, explicado novamente pelos valores pequenos do alcance ( $3\Phi$ ) igual a 1 m nas áreas FEI e FMA, e um alcance prático pouco acima do limite do *grid* amostral na FES (AP = 44 m).

TABELA 17 - SELEÇÃO DOS MODELOS COM AJUSTE POR MÉTODO DE COMPARTILHAMENTO DE PARÂMETROS NAS ÁREAS FEI, FES, FMA PELO CRITÉRIO DE INFORMAÇÃO DE AKAIKE (AIC)

Modelos	Parâmetros				L-V	Npar	(AP)	AIC
	$\beta$	$\tau^2$	$\sigma^2$	$\Phi$				
11*	1,1,2	-	1,2,2	1,1,2	-54,0	6	(82) (82) (10)	120,10
12	1,2,3	-	1,2,2	1,2,1	-53,9	7	(6) (47) (6)	120,11
09	1,1,1	-	1,1,1	1,1,1	-57,1	3	(1) (1) (1)	120,24
13	1,2,3	-	1,2,3	1,2,3	-52,9	9	(1) (46) (5)	123,83
10	1,1,1	-	1,1,1	1,1,2	-57,0	5	(44) (44) (8)	124,15
01	1,1,1	1,1,1	1,1,1	1,1,1	-58,4	4	(9) (9) (9)	124,99
04	1,1,2	1,1,2	1,2,3	1,2,2	-54,0	9	(2) (66) (66)	125,02
06	1,1,2	1,1,2	1,2,3	1,1,2	-54,0	9	(75) (75) (1)	125,11
02	1,2,3	1,1,2	1,1,1	1,1,1	-55,5	7	(16) (16) (16)	125,18
07	1,2,3	1,2,3	1,1,1	1,2,3	-52,9	10	(3) (47) (9)	125,86
05	1,1,2	1,2,3	1,2,2	1,2,2	-54,2	9	(8) (35) (35)	126,59
03	1,1,2	1,1,2	1,1,2	1,1,2	-56,7	8	(39) (39) (10)	129,55
08	1,2,3	1,2,3	1,2,3	1,2,3	-52,8	12	(1) (44) (1)	129,77

FONTE: O autor (2014).

NOTA: \* = melhor modelo pelo critério de AIC;  $\beta$  = vetor de médias;  $\tau^2$  = *nugget*;  $\sigma^2$  e  $\Phi$  = correlação espacial; L-V = log-verossimilhança; Npar = número de parâmetros; (AP) = alcance prático na ordem (FEI) (FES) (FMA) para  $3\Phi$  para  $Kappa = 0,5$ ; números “1,1,1; 1,1,2; 1,2,2; 1,2,1; 1,2,3” indicam a combinação das áreas na ordem “FEI, FES, FMA” em relação aos parâmetros.

Referente aos resultados gerais de desempenho dos modelos com ajuste individual e compartilhado sendo: Sem componente espacial (Nsm), por máxima verossimilhança (MV) (TABELA 16), e os modelos compartilhados sem *nugget* fixo e com *nugget* fixo (TABELA 17), foi possível relacionar os ganhos obtidos na comparação entre um ajuste e outro, pela seleção do critério do AIC, sendo nítido o escalonamento dos modelos em relação às técnicas de ajustes empregados.

Foram selecionados os modelos que tiveram os melhores resultados em relação ao valor do teste de AIC e grau de dependência espacial ( $\Phi$ ), nos ajustes pelos métodos compartilhados com e sem *nugget* fixo, de forma a observar se houve semelhança dos parâmetros compartilhados entre eles, reforçando a interpretação biológica nos modelos candidatos as estimativas e previsões de superfície.

Pode-se observar na Tabela 18 que a média ( $\beta$ ) convergiu com as mesmas características nos modelos, destacando que a maior dependência espacial se dá quando combinadas as áreas FEI-FES, para explicar os volumes médios ocorridos nas áreas de estudo (modelo 11). Os modelos “06” e “04” combinaram a mesma informação dos parâmetros ( $\beta$ ,  $\tau^2$ ,  $\sigma^2$ ), ocorrendo dependência espacial em até 66 m na FES e FMA no modelo “04” e até 75 m no modelo “06” para as áreas FEI e FES.

Esse resultado do ponto de vista geoestatístico é importante, pois, mostra a homogeneidade dos valores de volume entre as duas áreas distantes 3,5 km entre si, que combinaram força em relação aos seus parâmetros explicando a variabilidade espacial existente entre elas, com uma boa sobreposição do alcance prático (AP) em relação ao *grid* de amostragem empregado no estudo.

TABELA 18 - SIMILARIDADE DOS MODELOS DE COMPARTILHAMENTO DE PARÂMETROS COM MAIOR GRAU DE DEPENDÊNCIA ESPACIAL ( $\Phi$ ) NAS ÁREAS FEI, FES, FMA

Modelo	Parâmetros				Npar	(AP)	AIC
	$\beta$	$\tau^2$	$\sigma^2$	$\Phi$			
11*	1,1,2	-	1,2,2	1,1,2	6	(82) (82) (10)	120,10
04	1,1,2	1,1,2	1,2,3	1,2,2	9	(2) (66) (66)	125,02
06	1,1,2	1,1,2	1,2,3	1,1,2	9	(75) (75) (1)	125,11

FONTE: O autor (2014).

NOTA: \* = melhor modelo pelo critério de AIC;  $\beta$  = vetor de médias;  $\tau^2$  = *nugget*;  $\sigma^2$  e  $\Phi$  = correlação espacial; L-V = log-verossimilhança; Npar = número de parâmetros; (AP) = alcance prático na ordem (FEI) (FES) (FMA) para 3 $\Phi$  para *Kappa* = 0,5; números “1,1,1; 1,1,2; 1,2,2; 1,2,1; 1,2,3” indicam a combinação das áreas na ordem “FEI, FES, FMA” em relação aos parâmetros.

As tabelas 19, 20 e 21 evidenciam a separabilidade resultante dos efeitos fixos e aleatórios aplicados aos parâmetros dos modelos “11”, “06” e “04”, representadas utilizando a concepção dos diagramas de Venn.

TABELA 19 - SEPARABILIDADE DOS PARÂMETROS DO MODELO “11” EM RELAÇÃO AO GRAU DE DEPENDÊNCIA ESPACIAL ( $\Phi$ )

Parâmetros	Áreas	
$\beta$	$(FEI = FES) \neq FMA$	
$\tau^2$	fixo	
$\sigma^2$	$FEI \neq (FES = FMA)$	
$\Phi$	$(FEI = FES) \neq FMA$	

FONTE: O autor (2014).

NOTA:  $\beta$  = vetor de médias;  $\tau^2$  = *nugget*;  $\sigma^2$  e  $\Phi$  = parâmetros de correlação espacial; FEI, FES, FMA = ordem de sequência das áreas.

Na Tabela 19 observa-se que, provavelmente a composição das espécies no local em relação aos espaços ocupados resultou em uma distinção da variabilidade espacial ( $\sigma^2$ ) no fragmento (FEI), isso significa que, apesar das médias dos volumes e do alcance ( $\Phi$ ) serem semelhantes para as áreas FEI e FES, a maneira como elas se distribuem no espaço possui uma maior correlação nas áreas FES e FMA.

TABELA 20 - SEPARABILIDADE DOS PARÂMETROS DO MODELO “06” EM RELAÇÃO AO GRAU DE DEPENDÊNCIA ESPACIAL ( $\Phi$ )

Parâmetros	Áreas	
$\beta$	$(FEI = FES) \neq FMA$	
$\tau^2$	$(FEI = FES) \neq FMA$	
$\sigma^2$	$FEI \neq FES \neq FMA$	
$\Phi$	$(FEI = FES) \neq FMA$	

FONTE: O autor (2014).

NOTA:  $\beta$  = vetor de médias;  $\tau^2$  = *nugget*;  $\sigma^2$  e  $\Phi$  = parâmetros de correlação espacial; FEI, FES, FMA = ordem de sequência das áreas.



Pode-se notar nas Tabelas 20 e 21 que o parâmetro variabilidade espacial dos volumes por hectare foi distinto entre as áreas para os modelos “06” e “04”. Em análise destes, em relação ao modelo “11” possivelmente essa diferença ocorreu por se tratarem de modelos com mesmo nível de informação, diferente do modelo “11” que isolou o erro ( $\tau^2$ ) à nível de parcela, gerando com dessa forma maior precisão das estimativas.

TABELA 21 – SEPARABILIDADE DOS PARÂMETROS DO MODELO “04” EM RELAÇÃO AO GRAU DE DEPENDÊNCIA ESPACIAL ( $\Phi$ )

Parâmetros	Áreas	
$\beta$	(FEI = FES) $\neq$ FMA	
$\tau^2$	(FEI = FES) $\neq$ FMA	
$\sigma^2$	FEI $\neq$ FES $\neq$ FMA	
$\Phi$	(FEI = FES) $\neq$ FMA	

FONTE: O autor (2014).

NOTA:  $\beta$  = vetor de médias;  $\tau^2$  = *nugget*;  $\sigma^2$  e  $\Phi$  = parâmetros de correlação espacial; FEI, FES, FMA = ordem de sequência das áreas.

#### 5.4 KRIGAGEM E PREDIÇÃO DO VOLUME POR HECTARE

Fez-se a krigagem a partir dos modelos com os melhores resultados obtidos pelo critério de informação de Akaike (AIC) e grau de dependência espacial observado pelo parâmetro ( $\Phi$ ) em ambos os métodos de ajustes: parâmetros compartilhados sem *nugget* fixo “Modelo 04” e parâmetros compartilhados com *nugget* fixo “Modelo 11” e uma terceira krigagem nas áreas ajustadas em separado (não compartilhadas), a título de comparação entre os métodos, com representação das superfícies de interpolação estimadas para a variável volume em metros cúbicos por hectare nas áreas de estudo.

Não foi realizada a krigagem no modelo “06”, pois, o tipo de informação gerada visualmente na interpolação por este modelo é muito semelhante a representação do modelo “11”, com a desvantagem de uma menor dependência espacial.

#### 5.4.1 Krigagem e predição do modelo compartilhado sem *nugget* fixo

A krigagem realizada no modelo sem *nugget* fixo “04” mostra que ocorre a dependência espacial da variável volume entre pontos observados, na condição em que as médias ( $\beta$ ), o erro aleatório ( $\tau^2$ ) e a variabilidade espacial ( $\sigma^2$ ) são semelhantes nas áreas FEI e FES (FIGURAS 15A e 15B), mas, o grau de correlação espacial ( $\Phi$ ) interage com mais força entre os fragmentos FES e FMA (FIGURAS 15B e 15C).

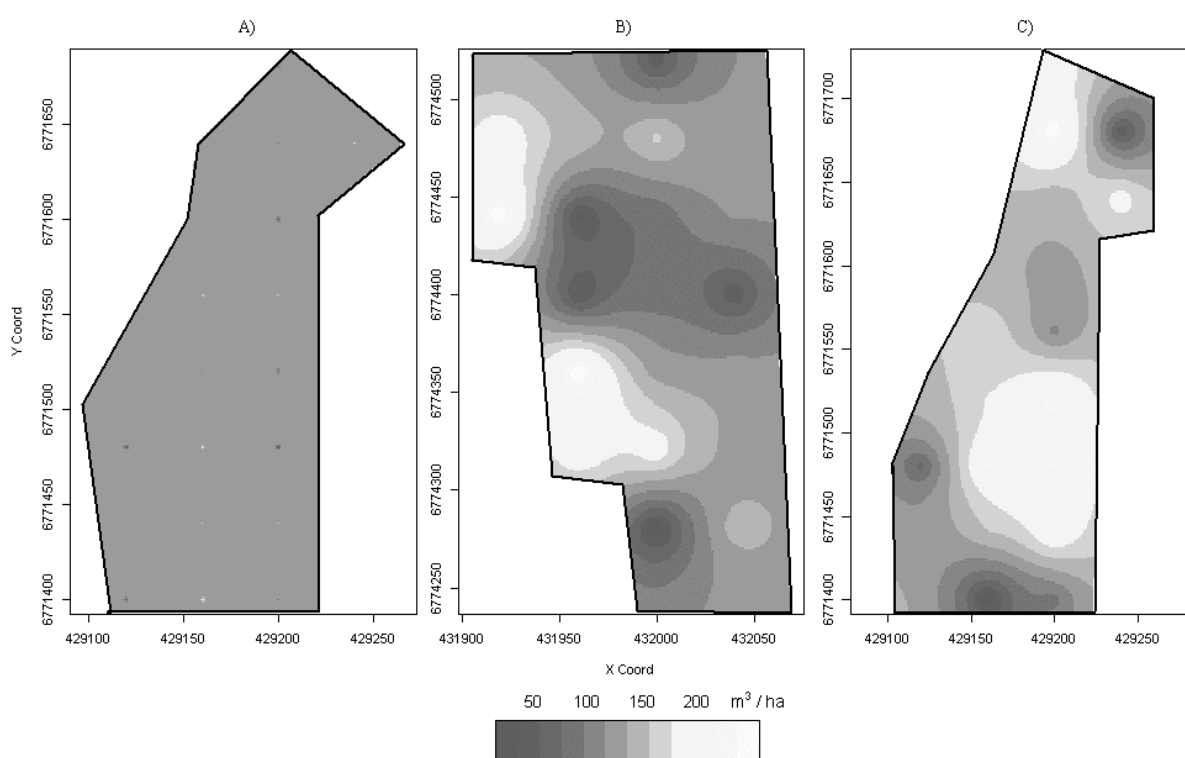


FIGURA 15 – PREDIÇÃO DOS VOLUMES ( $\text{m}^3 \text{ ha}^{-1}$ ) PARA O MODELO COMBINADO SEM *NUGGET* FIXO “04” NAS ÁREAS A) FEI, B) FES E C) FMA.  
FONTE: O autor (2014)

Observa-se na Figura 15 que os volumes distribuíram-se com uma amplitude de variação entre 40 a 250  $\text{m}^3 \text{ ha}^{-1}$  nas áreas de FES (FIGURA 15B) e FMA (FIGURA 15C), sendo que a maior parte dos volumes estimados ficou distribuída com valores abaixo de 170  $\text{m}^3 \text{ ha}^{-1}$ .

O pequeno alcance do modelo “04” ( $AP = 2 \text{ m}$ ) especificamente para a área de FEI, não foi suficiente para influenciar a média local, não gerando dessa forma

informação dos volumes por hectare nesse fragmento (FIGURA 15A).

#### 5.4.2 Krigagem e predição do modelo compartilhado com *nugget* fixo

A Figura 16 ilustra o resultado da krigagem realizada no modelo “11” com o *nugget* fixo ( $\tau^2$ ) e combinando os parâmetros de média ( $\beta$ ) e alcance ( $\Phi$ ) nas áreas FEI e FES, com uma variabilidade espacial ( $\sigma^2$ ) semelhantes nas áreas FES e FMA.

Observa-se que o modelo para as áreas de FEI e FES das Figuras 16A e 16B respectivamente, resultou em uma maior continuidade espacial devido o alcance prático obtido (AP = 82 m), bem como, uma suavização da superfície interpolada resultante de um maior número de pontos krigados, o que não foi constatado na área de FMA.

Na Figura 16C podemos notar que a média foi pouco influenciada pelo processo de interpolação, tendo a superfície predita gerado círculos com um alcance prático muito próximo ao centro das unidades amostrais (AP = 10 m), na área de Floresta Madura (FMA).

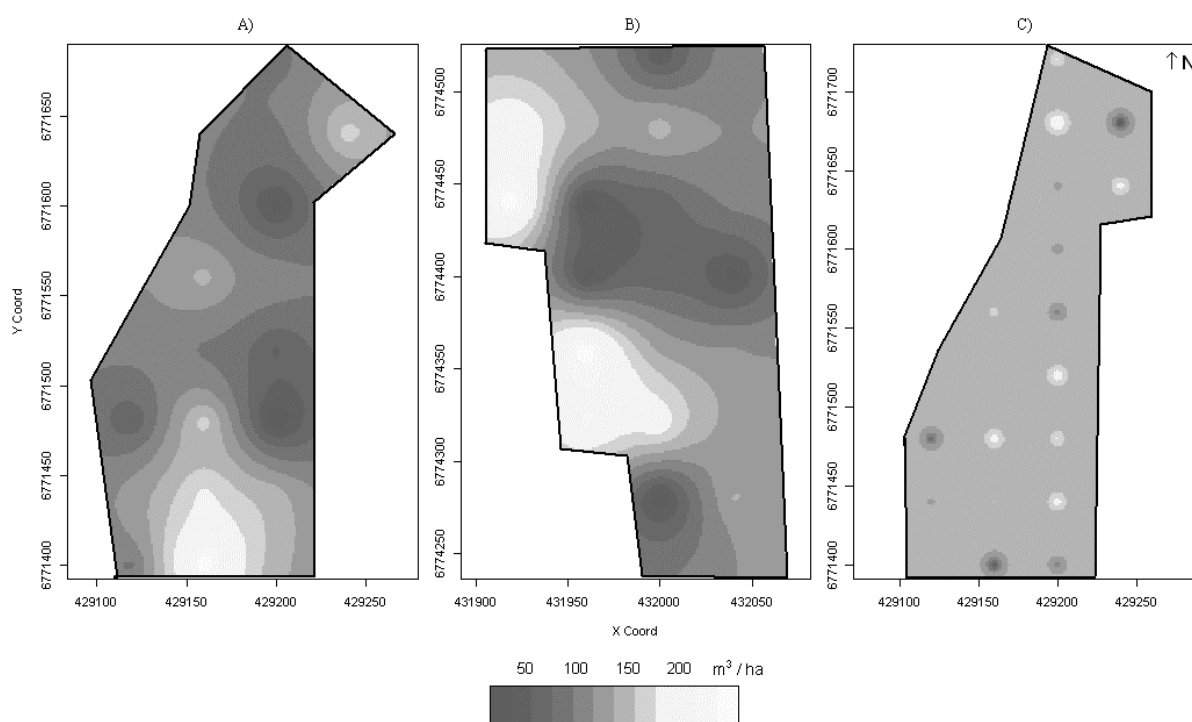


FIGURA 16 – PREDIÇÃO DOS VOLUMES ( $\text{m}^3 \text{ha}^{-1}$ ) PARA O MODELO COMPARTILHADO COM *NUGGET* FIXO “11” NAS ÁREAS A) FEI, B) FES E C) FMA.  
FONTE: O autor (2014)

### 5.4.3 Krigagem e predição do modelo não compartilhado

O resultado da krigagem do volume por hectare para o modelo ajustado por máxima verossimilhança (MV) efetuada nas áreas FEI, FES e FMA individualmente, teve um alcance prático de 8,30 m na FEI, 44,03 m na FES e 0,01 m na FMA, Figuras 17A, 17B e 17C respectivamente. Nota-se que na condição de ajuste proposto, não haveria continuidade espacial ou mesmo a informação das médias locais em duas das três áreas do estudo (FIGURA 17A e FIGURA 17C), representando parcialmente a relação espacial dos volumes nestes locais.

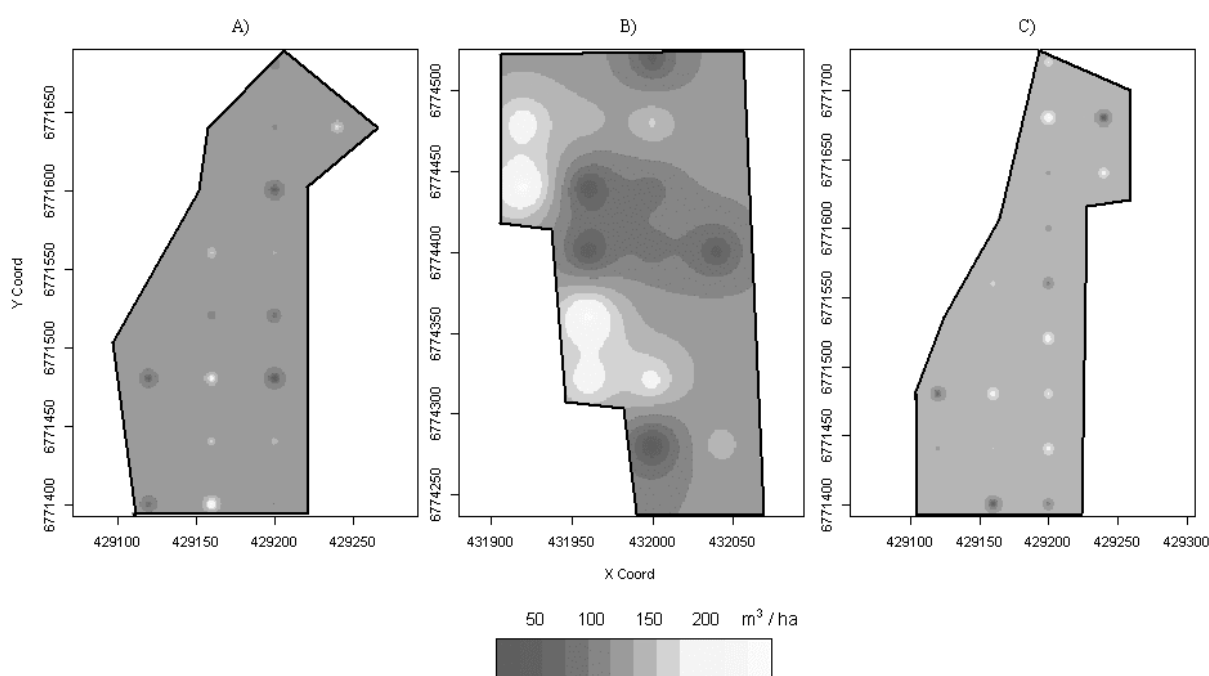


FIGURA 17 – PREDIÇÃO DOS VOLUMES ( $m^3 ha^{-1}$ ) PARA O MODELO NÃO COMPARTILHADO DE MÁXIMA VEROSSIMILHANÇA NAS ÁREAS A) FEI, B) FES E C) FMA.

FONTE: O autor (2014)

### 5.5 VALIDAÇÃO CRUZADA

Foi selecionado o modelo compartilhado com *nugget* fixo “11” conforme os valores obtidos pelos critérios de menor valor de Akaike (AIC) e maior alcance prático (AP) para validação da predição nas áreas, onde os resultados para o erro absoluto (EA) em  $m^3 ha^{-1}$ , desvio-padrão dos erros reduzidos ( $S_{ER}$ ) e erros médios reduzidos (ER) constam na Tabela 22.

TABELA 22 – RESULTADO DA VALIDAÇÃO CRUZADA PARA OS DADOS DE VOLUME UTILIZANDO O MODELO “11” NAS ÁREAS FEI E FES POR UNIDADE AMOSTRAL.

U.A.	FEI			FES		
	EA	S <sub>ER</sub>	ER	EA	S <sub>ER</sub>	ER
1	3,44	-0,6934	0,3738	0,58	-0,1171	0,1532
2	3,32	0,5924	0,3251	0,54	-1,2150	1,6468
3	3,41	-0,3744	0,2028	0,56	0,3167	0,4226
4	3,40	-0,2641	0,1433	0,56	0,6225	0,8316
5	3,38	-0,0734	0,0399	0,53	0,6491	0,8916
6	3,38	-0,1300	0,0707	0,55	-0,0427	0,0578
7	3,46	-0,8254	0,4440	0,55	1,2432	1,6838
8	3,35	0,2119	0,1158	0,53	-0,1185	0,1631
9	3,48	-1,0704	0,5734	0,55	0,1040	0,1409
10	3,41	-0,4510	0,2441	0,54	-0,9939	1,3510
11	3,43	-0,6122	0,3305	0,53	-0,1936	0,2664
12	3,37	-0,0117	0,0064	0,55	-0,7811	1,0574
13	3,39	-0,1867	0,1014	0,56	1,3229	1,7665
14	3,47	-0,9489	0,5093	0,53	-0,9968	1,3694
15	3,41	-0,3966	0,2148	0,53	-0,2783	0,3823
16	3,36	0,1307	0,0713	0,56	0,0978	0,1306
17	3,42	-0,4991	0,2699	0,57	0,6591	0,8760
18	--	--	--	0,54	0,1925	0,2615
19	--	--	--	0,54	0,5224	0,7083
20	--	--	--	0,58	-0,9385	1,2285
Totais	57,88	± 5,6023	--	10,98	± 0,0546	--
Médias	3,40	± 0,3295	--	0,55	± 0,0027	--

FONTE: O autor (2014)

NOTA: U.A = unidades amostrais; (EA) = erro absoluto dos volumes em m<sup>3</sup> ha<sup>-1</sup>; (S<sub>ER</sub>) = desvio-padrão dos erros reduzidos e (ER) = erros médios reduzidos.

Em relação aos resultados da validação cruzada (TABELA 22), observa-se que o modelo selecionado estimou de maneira satisfatória os volumes para as áreas, com um erro médio absoluto (EA) de 3,40 e 0,55 m<sup>3</sup>/ha para a FEI e FES respectivamente. As estimativas de erro absoluto (EA), desvio-padrão dos erros reduzidos (S<sub>ER</sub>) e erros médios reduzidos (ER) na área de FMA geraram extrapolações, tendo em vista que as médias na superfície interpolada pela krigagem nessa área foram pouco influenciadas, dado o baixo alcance prático atingido nesse fragmento (AP = 10 m).

Os totais acumulados das estimativas dos erros absolutos (EA) foram de 57,88 e 10,98 m<sup>3</sup>/ha para FEI e FES respectivamente, com os desvios-padrão dos erros reduzidos ( $S_{ER}$ ) total de  $\pm 5,6023$  para área FEI e  $\pm 0,0546$  m<sup>3</sup>/ha na FES, considerados valores ótimos para ambas as áreas, principalmente na FES, tendo as melhores estimativas em relação aos erros calculados, provavelmente influenciados por um maior número de amostras nessa área. Em relação aos erros médios reduzidos (ER), esse oscilou bastante entre os pontos observados nas unidades amostrais, ficando próximos a “1” em onze das vinte amostras apenas na FES.

Comparando-se os resultados obtidos pelo método de Prodan com a estimativa realizada pela krigagem utilizando o modelo de predição “11”, para os totais em volume por hectare nas áreas FEI e FES, nota-se que o estimador geoestatístico retornou valores muito próximos aos calculados pelo estimador clássico (Prodan) com uma diferença de 5,60 e 0,50 m<sup>3</sup>/ha para FEI e FES respectivamente.

Os resultados dos volumes totais obtidos pelo método estatístico foram 295,35 m<sup>3</sup>/ha na FEI e 435,08 m<sup>3</sup>/ha na FES, enquanto que os totais estimados pela krigagem foram 300,95 e 435,03 m<sup>3</sup>/ha respectivamente para FEI e FES.

TABELA 23 – COMPARAÇÃO ENTRE O ESTIMADOR CLÁSSICO (PRODAN) E O ESTIMADOR GEOESTATÍSTICO (KRIGAGEM) PARA O VOLUME TOTAL

Métodos	FEI	FES
	Volume (m <sup>3</sup> / ha)	
Estimador clássico (Prodan)	295,35	435,08
Estimador geoestatístico (Krigagem)	300,95	435,03
Diferença	5,60	0,50

FONTE: O autor (2014)

Os resultados corroboram com os métodos adotados para seleção do melhor modelo baseados na verossimilhança, a partir dos critérios analisados em conjunto como “AIC” e alcance prático “AP” mesmo em amostras pequenas.

Os estimadores consistentes da máxima verossimilhança garantem que os valores estimados estarão mais próximos dos valores reais observados, quanto maior for a amostra.

## 6 CONCLUSÕES

Os modelos que ajustaram as áreas de forma individual não foram satisfatórios no estudo, pois, a correlação espacial encontrada, bem como o alcance prático (AP) resultante foram pequenos.

Os modelos ajustados pelo método compartilhado com *nugget* fixo são superiores aos modelos sem *nugget* fixo segundo o critério de Akaike, sugerindo que o efeito do erro aleatório ( $\tau^2$ ) mensurado fora do semivariograma (intra-parcela) pode aumentar a correlação espacial dos volumes nas áreas.

O melhor ajuste foi obtido para o modelo “11” com “AIC = 120,10” é “AP = 82,50 m”, com forte correlação quando compartilhados o alcance ( $\Phi$ ) e as médias dos volumes ( $\beta$ ) nas áreas FEI e FES, combinados a uma variância espacial ( $\sigma^2$ ) semelhante para FES e FMA.

Comparando os modelos com parâmetros não compartilhados, os modelos compartilhados aumentaram a precisão das estimativas, pois, o maior valor do AIC dos modelos compartilhados (AIC = 129,77 m), é inferior ao menor valor encontrado nos modelos não compartilhados (AIC = 160,2 m).

O critério de Akaike (AIC) foi eficiente na comparação dos diferentes níveis de ajustes propostos, selecionando um modelo compatível com os padrões de distribuição espacial dos volumes nas áreas, sugere-se seu uso juntamente com a avaliação do AP em trabalhos utilizando *grid* de espaçamento regular.

Em relação à predição nas áreas, os modelos compartilhados mostram-se mais eficientes que os modelos ajustados individualmente, pois, é possível recuperar a informação dos volumes por hectare em duas das três áreas preditas.

A validação cruzada confirma a escolha do modelo “11” com valores baixos e próximos à zero para o erro absoluto (EA) e desvio-padrão dos erros reduzidos

( $S_{ER}$ ), bem como, um erro total acumulado de 57,88 e 10,98 m<sup>3</sup>/ha para FEI e FES respectivamente.

Devido a baixa influência da média na superfície interpolada da FMA do modelo “11”, a validação cruzada retornou valores extrapolados de erro absoluto (EA), desvio-padrão dos erros reduzidos ( $S_{ER}$ ) e erros médios reduzidos (ER).

Os volumes totais calculados pelo estimador geoestatístico e estimador clássico (Prodan) tiveram diferenças de 5,60 m<sup>3</sup>/ha na FEI e 0,50 m<sup>3</sup>/ha na FES, considerada pequena em relação ao tamanho da amostra.

Havendo necessidade de transformação dos dados para posterior aplicação dos modelos compartilhados, é importante que as áreas a serem concatenadas possuam o mesmo perfil de log-verossimilhança, ou seja, mesmo valor de lambda ( $\lambda$ ) garantindo a análise conjunta das mesmas.

Sugere-se o uso de modelos geoestatísticos com abordagem compartilhada “*Borrow Strength*” para estudos de variabilidade espacial em áreas fragmentadas, com tipologias semelhantes e presença de correlação espacial.

Em florestas naturais onde o arranjo errático da maioria das espécies pode acarretar em fraca dependência espacial, são indicados ajustes baseados na máxima verossimilhança.



## 7 RECOMENDAÇÕES

Deve ser enfatizado que as técnicas aplicadas no trabalho, apesar de conclusivas para o estudo, podem e devem ser ampliadas para outras situações, devido a flexibilidade da programação no qual foram construídas. Assim sendo, constitui-se uma iniciativa piloto, pois, a abordagem em relação a modelos mistos e modelos envolvendo parâmetros espaciais compartilhados, possuem grandes possibilidades de crescimento e adequações aplicadas às pesquisas na área do manejo e inventário florestal.

O estudo pode ser desdobrado na obtenção de classificadores de sítio, aproveitando a possibilidade de gerar médias móveis tanto em povoamentos florestais como em florestas nativas.

Há a necessidade de mais trabalhos referentes aos modelos mistos e compartilhados para predição de variáveis dendrométricas com enfoque espacial relacionados às florestas naturais, que se configuram um desafio maior na modelagem geoestatística.

## REFERÊNCIAS

- AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, Boston, v. 19, n. 6, p. 716–723, Dec. 1974.
- AMARAL, L. P.; MICHELLE DULLIUS, M.; GREFF, L. T. B., FERREIRA, R. A. DALMOLIN, R. S. D.; SCHNEIDER, P. R. Dependência espacial de diferentes alturas das árvores em um fragmento de floresta secundária na região do rebordo do Planalto do Rio Grande do Sul. In: **SIMPÓSIO LATINO AMERICANO DE MANEJO FLORESTAL**, 5., 2011, Santa Maria. Anais... Santa Maria: Universidade Federal de Santa Maria, 2011. p. 365-373.
- ANDRIOTTI, J. L. S., Introdução à Geoestatística. In: **Acta Geológica Leopoldensia**, São Leopoldo, v. 11, p. 5-81, 1988.
- BARBER, J. J.; GELFAND, A. E. Hierarchical spatial modeling for estimation of population size. **Environmental and Ecological Statistics**, New York, v. 14, p. 193-205, 2007.
- BAS, D.; VAN PUIJENBROEK, R. A. G. Conditional Regression Analysis. **Quality & Quantity**, Dordrecht, v. 23, issue 1, p. 83-108, 1989.
- BERG, M. V. D; KLAMT, E. Variabilidade espacial de característica de solos na região do Planalto Médio, RS: II. Análise da semivariância e da variância. **Revista Brasileira de Ciência do Solo**, Viçosa, v.21, n.3, 1997.
- BOX, G. E .P.; COX, D. R. An Analysis of Transformations. **Journal of the Royal Statistical Society**, London, v. 26, n. 2, p. 211-252, 1964.
- BOZDOGAN, H. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. **Psychometrika**, Williamsburg, v. 52, n. 3, p. 345-370, 1987.
- BONAT, W. H.; RIBEIRO JÚNIOR, P. J.; ZEVIANI, W. M. Comparando predições por modelos geoestatístico e aditivo generalizado para reconstituição de superfícies contínuas gaussianas. **Energia na Agricultura**, v. 26, n.2, p. 119-128, 2011.
- BUONGIORNO, J.; GILLESS, J. K. **Decision methods for forest resource management**. San Diego: Academic Press, 2003. 429p.
- BRILLINGER, D. **Selected works in probability and statistics**. Oslo: Springer, 1997. 609p.
- BURNHAM, P. K.; ANDERSON, D. R. **Model selection and multimodel inference: a practical information: theroretic approach**. 2<sup>nd</sup>. ed. Colorado: Springer, 2002. 347p.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística básica**. 6ª ed. São Paulo: Atual, 2009. 321p.

CALEGARIO, N.; MAESTRI, R.; LEAL, C. L.; DANIELS, R. Estimativa do crescimento de povoamento de Eucalyptus baseada na teoria dos modelos não lineares em multinível de efeito misto. **Ciência Florestal**, Santa Maria, v. 15, n. 3, p. 285-292, 2009.

CAMBARDELLA, C. A.; MOORMAN, T. B.; NOVAK, J. M.; PARKIN, T. B.; KARLEN, D. L.; TURCO, R. F.; KONOPKA, A. E. Field-scale variability of soil properties in Central Iowa Soils. **Soil Science Society of America Journal**, Madison, v. 58, n. 5, p. 1501-1511, Sept. /Oct. 1994.

CAMPBELL, D. T.; STANLEY, J. C. **Delineamentos experimentais e quase-experimentais de pesquisa**. São Paulo: USP/EPU, 1979. 138 p.

CARVALHO, T. M. **Variabilidade espacial de propriedades físico-hídricas de um Latossolo Vermelho Amarelo através da geoestatística**. 84f. Dissertação (Mestrado em Solos e Nutrição de Plantas) - Escola Superior de Agricultura de Lavras, Lavras, MG, 1991.

CHEN, J.; HOPMANS, J. W.; FOGG, G. E. Sampling design for soil moisture measurements in large field trials. **Soil Science**, New Brunswick. vol. 159, n. 3, p. 155-161, 1995.

COCHRAN, W. G. **Técnicas de Amostragem**. São Paulo: Ed. Fundo de Cultura, 1965. 539p.

COHEN, W. B.; SPIES, T. A.; BRADSHAW, G. A. Semivariograms of digital imagery for analysis of conifer canopy structure. **Remote Sensing of Environment**, New York, v. 34, n. 3, p. 167-178, 1990.

COSTA NETO, P. L. O. **Estatística**. São Paulo: Edgard Blucher, 1977. 264 p.

DEUTSCH, C. V.; JOURNEL, A. G. **GSLIB: Geostatistical software library and user's guide**. New York: Oxford University Press, 1992. 336p.

DEMIDENKO, E. **Mixed Models: theory and applications with R**. 2ª ed. Canada: Willey, 2013. 637p.

DIGGLE, P. J.; RIBEIRO JUNIOR, P. J. **Model-Based geostatistics**. New York: Springer, 2007. 228p.

DIGGLE, P. J.; TAWN, J. A.; MOYEED, R. A. Model-based geostatistics. **Journal of the Royal Statistical Society. Series C: applied statistics**, v. 47, n. 3, p. 299-350, 1998.

EGUCHI, E. S. **Variabilidade espacial de atributos físico-químicos de um solo hidromorfo no Município de Lavras, MG**. 85 f. Dissertação (Mestrado) – Universidade Federal de Lavras, Lavras, 2001.

EMILIANO, P. C. **Fundamentos e aplicações dos critérios de informação: Akaike e Bayesiano**. 92f. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras. Lavras, 2009.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA (EMBRAPA). Centro Nacional de Pesquisa de Solos. **Manual de métodos de análise de solo**. 2ª ed. Rio de Janeiro: Ministério da Agricultura e Abastecimento, 1997. 212 p.

FANHA, A. B. **Caracterização de reservatório através de técnicas estatísticas multivariadas e modelagem estocástica no campo de baixa do algodão, Bacia Potiguar**. 103f. Dissertação (Mestrado em Geoengenharia de Reservatórios) – Universidade Estadual de Campinas, Campinas, 1994.

FARIAS, A. M. L.; LAURENCEL, L. C. **Estatística Descritiva**. Niterói, 2008. Disponível em: <[http://www.professores.uff.br/anafarias/estdesc\\_2008.pdf](http://www.professores.uff.br/anafarias/estdesc_2008.pdf)>. Acesso em: 14 jun. 2012.

FINLEY, A. O.; BANERJEE, S.; MACFARLANE, D. W. **A hierarchical model for quantifying forest variables over large heterogeneous landscapes with uncertain forest areas**. Michigan: Department of Forestry, 2010. 88p.

GREGOIRE, T. G.; SCHABENBERGER, O. Fitting bole-volume equations to spatially correlated within-tree data. In: ANNUAL CONFERENCE OF APPLIED STATISTICS IN AGRICULTURE, 6., 1994, Manhattan. **Proceedings...** Kansas, 1994. p. 120-133.

GUIMARÃES, E. C. **Variabilidade espacial da umidade e da densidade do solo em um Latossolo Roxo**. 135 f. Dissertação (Mestrado em Engenharia Agrícola) - Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas, 1993.

HARRELL, JR.; FRANK, E. **Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis**. New York: Springer, 2002. 572 p. (Spring Series in Statistics).

HERNANDEZ, J.; EMERY, X. A geostatistical approach to optimize sampling designs for local forest inventories. **Canadian Journal of Forest Research**, v. 39, p. 1465–1474, 2009.

HOAGLIN, D. C. Letras-resumo: um conjunto de estatísticas ordinais selecionadas. In: HOAGLIN, D. C.; MOSTELLER, F.; TUKEY, J.W. (Ed.). **Análise exploratória de dados. Técnicas robustas: um guia**. Lisboa: Salamandra, 1992. cap. 2, p. 32-58.

HUIJBREGTS, C. J. Regionalized variables and quantitative analysis of spatial data. In: DAVIS, J. C.; McCULLAGH, M. J. (Ed.). **Display and analysis of spatial data**. New York: John Wiley, 1975. p.38-53.

ISAAKS, E. H.; SRIVASTAVA, R. M. **Applied Geostatistics**. New York: Oxford University Press, 1989.

JELIAZKOV, I.; LEE, E. H. MCMC Perspectives on Simulated Likelihood Estimation. **In: Advances in Econometrics: Maximum Simulated Likelihood Methods and Applications**, 2010. v. 26, p. 3-39.

JOURNEL, A. G. **Geoestatistic for environmental science**. Las Vegas: Enviromental Protection Agency, 1988.

KANGAS, A.; MALTANO, M. **Forest Inventory: Methodology and Applications**. New York: Springer, 2006. 350p.

KENT, B. M.; DRESS, P. E. On the convergence of forest stand spatial pattern over time: the case of random initial spatial pattern. **Forest Science**, Bethesda, v. 25, n. 3, p. 445-451, 1979.

KLEIN, R. M. Aspectos fitofisionômicos da floresta estacional na falda da Serra Geral (RS). In: CONGRESSO NACIONAL DE BOTÂNICA, n. 34, 1983, Porto Alegre. **Anais...** Porto Alegre, 1983. v. 1, p. 73-110.

KITANIDIS, J. K. **Introduction to geostatistics: applications in hydrology**. New York: Cambridge University Press, 1997. 247p.

LANDIM, P. M. B. **Análise Estatística de Dados Geológicos**. São Paulo: Ed. UNESP, 1998. p. 225.

LONGHI, S. J. (Coord.). **Dinâmica de Três Fases Sucessionais de uma Floresta Estacional Decidual**. Santa Maria: UFSM, 22 de agosto de 2001 (Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul, n. 01/0673.8). Projeto concluído.

MACHADO, S. A.; FIGUEIREDO FILHO, A. **Dendrometria**. Curitiba: Os autores, 2003. 309p.

McBRATNEY, A.; WEBSTER, R. Choosing functions for semi-variograms os soil properties and fitting them to sample estimates. **Journal of Soil Science**, n. 37, p. 617-639, 1986.

MALCZEWSKI, J. **GIS and Multicriteria Decision Analysis**. New York: Wiley, 1999. 298p.

MATÈRN, B. **Spatial Variation**. Berlin: Springer, 1986. 594p.

MATHERON, G. The theory of regionalized variables and its applications. **Les Cahiers du Centre de Morphologie mathématique**, Fontainebleau, n. 5, cap. 3, p. 115-13, 1971.

MATTHIOPOUOS, J. **How to be a quantitative ecologist: the 'A to R' of green mathematics and statistics**. Scotland: Wiley, 2011. 1091p.

MELLO, J. M.; SCOLFORO, J. R. S. Análise comparativa de procedimentos de amostragem em um remanescente de Floresta Estacional Semidecídua Montana. **Revista Árvore**, Viçosa, v.24, n.1, p. 55-62, 2000.

MELLO, J. M. **Geoestatística aplicada ao inventário florestal**. 110f. Tese (Doutorado em Recursos Florestais) - Escola Superior de Agricultura Luiz de Queiroz. Piracicaba, 2004.

MILLAR, R. B. Simulated maximum likelihood applied to non-Gaussian and nonlinear mixed effect and state-space models. **Australian and New Zealand Journal of Statistics**, v. 46, issue 4, p. 543-544, 2004.

MORENO, J. A. **Clima do Rio Grande do Sul**. Porto Alegre: Secretaria da Agricultura, 1961. 73 p.

MÜLLER, H. G.; STADMÜLLER, U.; TABNAK, F. Spatial smoothing of geographically aggregated data, with application to the construction of incidence maps. **Journal of the American Statistical Association**, 92: 61-71, 1997.

NANOS, N.; CALAMA, R.; MONTERO, G.; GIL, L. Geostatistical prediction of height-diameter models. **Forest Ecology and Management**, n. 195, p. 221-235. 2004

OLEA, R. A. **Geostatistics for engineers and earth scientists**. New York: Springer, 1999. 303 p.

OLIVEIRA, M. C. N. **Métodos de estimação de parâmetros em modelos geoestatísticos com diferentes estruturas de covariância: uma aplicação ao teor de cálcio no solo**. 140f. Tese (Doutorado) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo. Piracicaba, 2003.

OLIVEIRA, M. S. **Planos amostrais para variáveis espaciais utilizando geoestatística**. Campinas. 100f. Dissertação (Mestrado) – Universidade Estadual de Campinas, 1991.

OLIVER, M. A.; WEBSTER, R. Kriging: a method of interpolation for geographical information systems. **International Journal of Geographical Information Systems**, v. 4, issue 3, p. 313-332, 1990.

ORTIZ, G. C. **Aplicação de métodos geoestatísticos para identificar a magnitude e a estrutura da variabilidade espacial de variáveis físicas do solo**. 75f. Dissertação (Mestrado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2002.

PEBESMA, E. J.; BIVAND, R. S. Classes and methods for spatial data in R. **R News**, v. 5, n. 2, 2005. Disponível em: <<http://cran.r-project.org/doc/Rnews/>>. Acesso em: 18 mar. 2014.

PÉLLICO NETTO, S.; BRENA, D. A. **Inventário Florestal**. Curitiba: Os autores, 1997. 316 p.

PÉLISSIER, R.; GOREAUD, F. A practical approach to the study of spatial structure in simple cases of heterogeneous vegetation. **Journal of Vegetation Science**, v.12, issue 1, p. 99-108, 2001.

PINHEIRO, J.; BATES, D. **Mixed-Effects Models in S and S-PLUS**. New York: Springer, 2000. 530p.

PLOMER, A.; DUTTER, R. New directions in geostatistics. **Journal of Statistical Planning and Inference**, Amsterdam, v. 91, p. 499-509, 2000.

RAMBO, B. A imigração da selva higrófila no Rio Grande do Sul. **Anais Botânicos do Herbário Barbosa Rodrigues**. Itajaí, v.3, n.3, p.55-91, 1951.

RAMON, C. L.; MILLIKEN, G. A.; STROUP, W. W.; SCHABENBERGER, O. **SAS for Mixed Models**. North Carolina: SAS institute, 2006. 795 p.

R CORE TEAM. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria: R Foundation for Statistical Computing, 2013. Disponível em: <<http://www.R-project.org/>>. Acesso em: 12 abr. 2014.

REITZ, R.; KLEIN, R. M.; REIS, A. **Projeto madeira do Rio Grande do Sul**. Porto Alegre: CORAG, 1988. 525 p.

RIBEIRO JUNIOR, P. J.; DIGGLE, P. J. The geoR package functions for geostatistical data analysis: **R. News**, v. 1, n. 2, p. 15-18, 2001.

RIBEIRO JÚNIOR, P. J. **Métodos geoestatísticos no estudo da variabilidade espacial dos parâmetros do solo**. 99 f. Dissertação (Mestrado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1995.

RODE, R.; FIGUEIREDO FILHO, A.; MACHADO, S. A.; GALVÃO, F. Análise do padrão espacial de espécies e de grupos florísticos estabelecidos em um povoamento de *Araucária angustifolia* e em uma Floresta Ombrófila Mista no Centro-Sul do Paraná. **Revista Floresta**, Curitiba, v. 40, n. 2, p. 255-268, 2010.

ROHLF, F. J.; SOKAL, R. R. **Statistical Tables**. San Francisco: Freeman & Comp, 1969. 641p.

RONGXIA, L.; PETE, B.; KEVIN, B. Informed Development of Meta Heuristics for Spatial Forest Planning Problems. **The Open Operational Research Journal**, n. 4, p. 1-11, 2010. Disponível em: <<http://www.benthamscience.com/open/toorj/>>. Acesso em: 20 mar. 2014.

SAMPER, F. J. S.; RAMIREZ, J. C. **Geoestadística: aplicaciones a la hidrogeologia subterránea**. Barcelona: Gráficas Torres, 1990. 484p.

SCHNEIDER, P. R.; FINGER, C. A. G. **Manejo Sustentado de Florestas Inequiâneas Heterogêneas**. Santa Maria: UFSM, 2000. 195p.

SCOLFORO, J. R. S.; MELLO, J. M. M. **Inventário Florestal**. Lavras: UFLA, 2006. 561p.

SOARES, A. G. **Geoestatística para as ciências da terra e do ambiente**. Lisboa: IST Press, 2000. 206p.

STRECK, E. V.; KAMPS, N.; DALMOLIN, R. S. D. **Solos do Rio Grande do Sul**. POA; Editora UFRGS, 2002. 107p.

TEDESCO, M. J.; GIANELLO, C.; BISSANI, C. A.; BOHNEN, H.; VOLKWEISS, S. J. **Análise de solo, plantas e outros materiais**. Porto Alegre: UFRGS / Departamento de Solos, 1995. 118p. (Boletim Técnico).

TRANGMAR, B. B.; YOST, R. S.; UEHARA, G. Application of geostatistics to spatial studies of soil properties. **Advances in Agronomy**, San Diego, CA, v.38, p. 45-93, 1985.

UNIVERSIDADE FEDERAL DE SANTA MARIA. **Inventário florestal contínuo do Rio Grande do Sul**. Santa Maria: Convênio UFSM/SEMA, 2001. 706 p. (Relatório final original).

VACCARO, S. **Caracterização fitossociológica de três fases sucessionais de uma Floresta Estacional Decidual, no município de Santa Tereza, RS**. 92 f. Dissertação (Mestrado em Engenharia Florestal) □ Centro de Ciências Florestais, UFSM, Santa Maria, 1997.

VENABLES, W. N.; RIPLEY, B. D. **Modern Applied Statistics with S**. 4th ed. New York: Springer, 2002.

VENDRUSCULO, L. G. **Uso de índice de desempenho e do critério de Akaike para ajuste de modelos de semivariograma**. Campinas: Embrapa Informática Agropecuária, 2003. 5p. (Comunicado técnico, 58).

VELOSO, H. P.; GOES FILHO, L. **Fitogeografia brasileira**: classificação fisionômica ecológica da vegetação neotropical. Brasília: MME, 1982. 80p. (Projeto RADAM Brasil. Série Vegetação).

VIEIRA, S. R. Geoestatística em estudos de variabilidade espacial do solo. In: NOVAES, R.S.; ALVAREZ V, V.H.; SCHAESER, C.E.G.R. (Ed.). **Tópicos em Ciências do Solo**. Viçosa: Sociedade Brasileira de Ciência do Solo, 2000. cap.1, p.1-54.

VIEIRA, S. R.; LOMBARDI NETO, F. Variabilidade espacial do potencial de erosão das chuvas do Estado de São Paulo. **Bragantia**, Campinas, v. 54, n. 2, p. 405-412, 1995.

VUOLO, J. H. **Fundamentos da Teoria de Erros**. 2. ed. São Paulo: Edgar Blücher, 1996. 247p.



ZAWADZKI, J.; CIESZEWSKI, C. J.; ZASADA, M.; LOWE, R. C. Applying geostatistics for investigations of forest ecosystems using remote sensing imagery. **Silva Fennica**, 2005. 39 n. 4, p. 599–617, 2005.

WHITMORE, T. C. **An introduction to Tropical Rain Forests**. New York: Oxford University Press, 1990. 226p

WOJCIECHOWSKI, J. C. **Geoestatística aplicada ao estudo das características físico-químicas do solo em área de Floresta Estacional Decidual**. 88 f. Dissertação (Mestrado em Geomática) – UFSM, Santa Maria, 2006.

WU, L. **Mixed Effects Models for Complex Data**. Canada: Tylor & Francis Group, 2010. 419p.

YAMAMOTO, J. K.; LANDIM, P. M. **Geoestatística: conceitos e aplicações**. São Paulo: Oficina de Textos, 2013. 215p.

**ANEXOS (Códigos em R)**

1 ANÁLISES ESTATÍSTICAS E EXPLORATÓRIAS.....	99
2 ESTIMATIVAS POR MÁXIMA VEROSSIMILHANÇA .....	101
3 KRIGAGEM E PREDIÇÃO.....	103
4 VALIDAÇÃO DO MODELO.....	108
5 FUNÇÕES DOS MODELOS COMPARTILHADOS .....	108

## 1 ANÁLISES ESTATÍSTICAS E EXPLORATÓRIAS

```
#####
## Autoria: Naimara Vieira do Prado e Paulo Justiniano Ribeiro Jr.    ##
## Modificado por: Julio Cesar Wojciechowski                          ##
#####

## Carregando pacotes
## Install.packages("geoR", "MASS", dep=TRUE)
## require(geoR)
## require(MASS)

## getwd()

## Carregando dados
##CAv<- read.geodata("CA_CURTO3.txt", head=TRUE, coords.col=1:2, data.col=11,
units.m.col=3)
##FSv <- read.geodata("FS_CURTO3.txt", head=TRUE, coords.col=1:2, data.col=11,
units.m.col=3)
##MMv<-read.geodata("MM_CURTO3.txt", head=TRUE, coords.col=1:2,data.col=11,
units.m.col=3)

## Gráfico exploratório dos dados
## plot(CAv)
## plot(FSv)
## plot(MMv)

## Grid de amostragem
## par(mfrow=c(1,3))
## boxplot(CAv$data, main="FEI",ylab="Vol / ha")
## boxplot(FSv$data, main="FES",ylab="Vol / ha")
## boxplot(MMv$data, main="FMA",ylab="Vol / ha")
```

```

# Valores de lambda - transformação Box-cox
## par(mfrow=c(1,3))
## CAvbc<-boxcox(CAv)
## FSvbc<-boxcox(FSv)
## MMvbc<-boxcox(MMv)

## Estatísticas descritivas dados originais
## summary(CAv$data); var(CAv$data)
## summary(FSv$data); var(FSv$data)
## summary(MMv$data); var(MMv$data)

## Transformando os dados (lambda = 0)

## CAv.t <- log(CAv$data); summary(CAv.t); var(CAv.t)
## FSv.t <- log(FSv$data); summary(FSv.t); var(FSv.t)
## MMv.t <- log(MMv$data); summary(MMv.t); var(MMv.t)

### Graficos de QQ-plot de percentagens acumuladas

## par(mfrow=c(1,3))
## qqnorm(CAv.t, main = "FEI"); qqline(CAv.t)
## qqnorm(FSv.t, main = "FES"); qqline(FSv.t)
## qqnorm(MMv.t, main = "FMA"); qqline(MMv.t)

## Histograma dados transformados + boxplot dos dados transformados

## par(mfrow=c(3,2))
## hist(CAv.t, prob=TRUE); lines(density(CAv.t))
## boxplot(CAv.t, main="FEI",ylab="Vol / ha")
## hist(FSv.t, prob=TRUE); lines(density(FSv.t))
## boxplot(FSv$data, main="FES",ylab="Vol / ha")
## hist(MMv.t, prob=TRUE); lines(density(MMv.t))
## boxplot(MMv$data, main="FMA",ylab="Vol / ha")

```

## 2 ESTIMATIVAS POR MÁXIMA VEROSSIMILHANÇA

```
## rm(list=ls())
## source("functions.R")
## Inserindo dados - Volume
## CAv<-read.geodata("CA_CURTO3.txt", head=TRUE, coords.col=1:2, data.col=11,
units.m.col=3)
##FSv <-read.geodata("FS_CURTO3.txt", head=TRUE, coords.col=1:2, data.col=11,
units.m.col=3)
##MMv<-read.geodata("MM_CURTO3.txt", head=TRUE, coords.col=1:2,data.col=11,
units.m.col=3)

## CAv$data <- log(CAv$data)
## FSv$data <- log(FSv$data)
## MMv$data <- log(MMv$data)

## Lista de dados (3 áreas)
## dados.3a <- list(CAv, FSv, MMv)
## fit <- list()

## fit$fit01 <- geoAreas(mu=c(1,1,1), sigma2=c(1,1,1), phi=c(1,1,1),
##
tau2=c(1,1,1), data=dados.3a, control=list(maxit=1000),
method="BFGS")

## fit$fit02 <- geoAreas(mu=c(1,2,3), sigma2=c(1,1,1), phi=c(1,1,1),
##
tau2=c(1,1,2), data=dados.3a, control=list(maxit=1000),
method="BFGS")

## fit$fit03 <- geoAreas(mu=c(1,1,2), sigma2=c(1,1,2), phi=c(1,1,2),
##
tau2=c(1,1,2), data=dados.3a, control=list(maxit=1000),
method="BFGS")

## fit$fit04 <- geoAreas(mu=c(1,1,2), sigma2=c(1,2,3), phi=c(1,2,2),
```

```

##                                tau2=c(1,1,2), data=dados.3a, control=list(maxit=1000),
method="BFGS")

## fit$fit05 <- geoAreas(mu=c(1,1,2), sigma2=c(1,2,2), phi=c(1,2,2),
##                                tau2=c(1,2,3), data=dados.3a, control=list(maxit=1000),
method="BFGS")

## fit$fit06 <- geoAreas(mu=c(1,1,2), sigma2=c(1,2,3), phi=c(1,1,2),
##                                tau2=c(1,1,2), data=dados.3a, control=list(maxit=1000),
method="BFGS")

## fit$fit07 <- geoAreas(mu=c(1,2,3), sigma2=c(1,1,1), phi=c(1,2,3),
##                                tau2=c(1,2,3), data=dados.3a, control=list(maxit=1000),
method="BFGS")

## fit$fit08 <- geoAreas(mu=c(1,2,3), sigma2=c(1,2,3), phi=c(1,2,3),
##                                tau2=c(1,2,3), data=dados.3a, control=list(maxit=1000),
method="BFGS")

## fit$fit00 <- geoAreas(mu=c(1,1,1), sigma2=c(1,1,1), phi=c(1,1,1),
##                                tau2=c(1,1,2), data=dados.3a, control=list(maxit=1000),
method="BFGS")

## Obtendo o nugget – Exemplo para área CAV
## Anova - CA longo
## Ca <- read.table("AREA1_longo.txt", head=TRUE)
## attach(Ca)
## Ca.geo <- as.geodata(Ca, coords.col=1:2, data.col=8)
## hist(log(Vol.ha), prob=TRUE); lines(density(log(Vol.ha)))
## Ca[,3] <- as.factor(Ca[,3])
## model <- aov(log(Vol.ha) ~ Pto, data=Ca)
## anova(model)
## tau1.est <- anova(model)$Mean[2]
## Ajuste incluindo o nugget

```

```
## fit$fix.nug09 <- geoAreas(mu=c(1,1,1), sigma2=c(1,1,1), phi=c(1,1,1),
##   tau2=c(tau1.est,tau2.est,tau3.est)/10,   fix.tau2=TRUE,   data=dados.3a,
control=list(maxit=1000), method="BFGS")
```

```
## fit$fix.nug10 <- geoAreas(mu=c(1,1,2), sigma2=c(1,1,1), phi=c(1,1,2),
##   tau2=c(tau1.est,tau2.est,tau3.est)/10,   fix.tau2=TRUE,   data=dados.3a,
control=list(maxit=1000), method="BFGS")
```

```
## fit$fix.nug11 <- geoAreas(mu=c(1,1,2), sigma2=c(1,2,2), phi=c(1,1,2),
##   tau2=c(tau1.est,tau2.est,tau3.est)/10,   fix.tau2=TRUE,   data=dados.3a,
control=list(maxit=1000), method="BFGS")
```

```
## fit$fix.nug12 <- geoAreas(mu=1:3, sigma2=c(1,2,2), phi=c(1,2,1),
##   tau2=c(tau1.est,tau2.est,tau3.est)/10, fix.tau2=TRUE,   data=dados.3a,
control=list(maxit=1000), method="BFGS")
```

```
## fit$fix.nug13 <- geoAreas(mu=c(1,2,3), sigma2=c(1,2,3), phi=c(1,2,3),
##   tau2=c(tau1.est,tau2.est,tau3.est)/10,   fix.tau2=TRUE,   data=dados.3a,
control=list(maxit=1000), method="BFGS")
```

```
## Resumo
```

```
## fit.res <- data.frame(logLik = sapply(fit, logLik), npar = sapply(fit, function(x)
x$npar), AIC = sapply(fit, AIC), BIC = sapply(fit, ##   BIC), conver = sapply(fit,
function(x) x$conver))
```

### 3 KRIGAGEM E PREDIÇÃO

```
## par(mfrow=c(1,1))
```

```
## Área 1 - CAv
```

```
## Construção da borda
```

```
## points(CAv)
```

```
## borda<-locator(type="l")
```

```
## borda1 <- read.table("bordaCA.txt", head=TRUE)
```

```
## polygon(borda1,col=4)
```

```

## lines(borda1)

## apply(borda1,2,range) #Mostra o mínimo e máximo das coordenadas
## Grid de interpolação
## gr1 <- expand.grid(x=seq(429096.8, 429265.9,by=1), y=seq(6771393, 6771689,
by=1))
## plot(gr1)

## Grid de interpolação com recorte da borda
## gi <- polygrid(gr1,bor=borda1)
## points(gi, pch="+", col=2)

## krigagem (lambda = 1 st e lambda = 0 log)
## kc1 <- krige.control(obj.model = NULL, beta = model.MPar[1,1], cov.model =
"matern", cov.pars =
## c(model.MPar[1,2],model.MPar[1,3]), kappa = 0.5, lambda=0, nugget =
model.MPar[1,4])
## d.k1 = krige.conv(CAv, loc=gr1, krige=kc1)
## krige.conv: model with constant mean
## krige.conv: performing the Box-Cox data transformation
## krige.conv: back-transforming the predicted mean and variance
## krige.conv: Kriging performed using global neighbourhood

## Mapa
## image(d.k1, loc=gr1, border=borda1, x.leg=c(429050, 429155), y.leg=c(6771655,
6771675), ##zlim=range(d.k1$predict))
## title(main = "FEI - vol/ha ", font.main=4)
## text(429310,6771680,expression(bold(NA%up%N)))

## Área 2 - FSv
## Construção da borda
## points(FSv)
## borda<-locator(type="l")

```



```

## borda2 <- read.table("bordaFS.txt", head=TRUE)
## polygon(borda2, col=4)
## lines(borda2)
## apply(borda2,2,range) #Mostra o mínimo e máximo das coordenadas
## Grid de interpolação
## gr2 <- expand.grid(x=seq(431905.4, 432069.0, by=1), y=seq(6774237, 6774525,
by=1))
## plot(gr2)

## Grid de interpolação com recorte da borda
## gi <- polygrid(gr2, bor=borda2)
## points(gi, pch="+", col=2)

## kc2 <- krige.control(obj.model = NULL, beta = model.MPar[2,1], cov.model =
"matern", cov.pars =
## c(model.MPar[2,2],model.MPar[2,3]), kappa =0.5, lambda=0, nugget =
model.MPar[2,4])
## d.k2 = krige.conv(FSv, loc=gr2, krige=kc2)

## Mapa
## image(d.k2, loc=gr2, border=borda2, x.leg=c(431850, 431975), y.leg=c(6774250,
6774270), zlim=range(d.k2$predict))
## title(main = "FES - vol/ha ", font.main=4)
## text(432110, 6774510, expression(bold(NA%up%N)))

## Área 3
## Construindo a borda
## points(MMv)
## borda<-locator(type="l")

## polygon(borda,col=4)
## borda<-data.frame(borda)
## borda<- rbind(borda,borda[1,])

```

```

## borda3 <- read.table("bordaMM.txt", head=TRUE)
## polygon(borda3,col=4)
## lines(borda3)

## apply(borda3,2,range) #Mostra o mínimo e máximo das coordenadas
## Grid de interpolação
## gr3 <- expand.grid(x=seq(429103.1, 429259.4, by=1), y=seq(6771392, 6771729,
by=1))
##plot(gr3)

## Grid de interpolação com recorte da borda
## gi <- polygrid(gr3, bor=borda3)
## points(gi, pch="+", col=2)

## krigagem
## kc3 <- krige.control(obj.model = NULL, beta = model.MPar[3,1], cov.model =
"matern", cov.pars =
## c(model.MPar[3,2],model.MPar[3,3]), kappa = 0.5, lambda=0, nugget =
model.MPar[3,4])

## d.k3 = krige.conv(MMv, loc=gr3, krige=kc3)
## image(d.k3, loc=gr3, border=borda3, x.leg=c(429030, 429150), y.leg=c(6771680,
6771700), zlim=range(d.k3$predict))
## title(main = "FMA - vol/ha ", font.main=4)
## text(429315, 6771710, expression(bold(NA%up%N)))

#### 3 mapas juntos
## par(mfrow=c(1,3))
## image(d.k1, loc=gr1, border=borda1, x.leg=c(429050, 429155), y.leg=c(6771655,
6771675), ##zlim=range(d.k1$predict))

## title(main = "FEI - vol/ha ", font.main=4)
## text(429310,6771680,expression(bold(NA%up%N)))

```

```
## image(d.k2, loc=gr2, border=borda2, x.leg=c(431850, 431975), y.leg=c(6774250,
6774270), ##zlim=range(d.k2$predict))
## title(main = "FES - vol/ha ", font.main=4)
## text(432110, 6774510, expression(bold(NA%up%N)))
## image(d.k3, loc=gr3, border=borda3, x.leg=c(429030, 429150), y.leg=c(6771680,
6771700), ##zlim=range(d.k3$predict))
## title(main = "FMA - vol/ha ", font.main=4)
## text(429315, 6771710, expression(bold(NA%up%N)))
```

```
##Análise para cada área individual
```

```
## CAv.ml <- likfit(CAv, lambda=0, ini=c(0.015, 40))
## FSv.ml <- likfit(FSv, lambda=0, ini=c(0.06, 40))
## MMv.ml <- likfit(MMv, lambda=0, ini=c(0.15, 40))
## CAv.ml <- likfit(CAv, ini=c(0.015, 40))
## FSv.ml <- likfit(FSv, ini=c(0.06, 40))
## MMv.ml <- likfit(MMv, ini=c(0.15, 40))
```

```
### Modelo espacial - 4 parâmetros é o modelo não espacial - 2 parâmetros.
```

```
summary(CAv.ml)
summary(FSv.ml)
summary(MMv.ml)
```

```
par(mfrow=c(1,3))
```

```
## Mapa área 1
```

```
## borda1 <- read.table("bordaCA.txt", head=TRUE)
## gr1 <- expand.grid(x=seq(429096.8, 429265.9,by=1), y=seq(6771393, 6771689,
## by=1))
## gi <- polygrid(gr1, bor=borda1)
## kc1 <- krige.control(obj=CAv.ml)
## d.k1 = krige.conv(CAv, loc=gr1, krige=kc1)
## image(d.k1, loc=gr1, border=borda1)
## title(main = "FEI - vol/ha ", font.main=4)
```

```
## Mapa área 2
## borda2 <- read.table("bordaFS.txt", head=TRUE)
## gr2 <- expand.grid(x=seq(431905.4, 432069.0, by=1), y=seq(6774237, 6774525,
## by=1))
## gi<- polygrid(gr2, bor=borda2)
## kc2 <- krige.control(obj= FSv.ml)
## d.k2 = krige.conv(FSv, loc=gr2, krige=kc2)
## image(d.k2, loc=gr2, border=borda2)
## title(main = "FES - vol/ha ", font.main=4)
```

```
## Mapa área 3
## Grid de interpolação
## borda3 <- read.table("bordaMM.txt", head=TRUE)
## gr3 <- expand.grid(x=seq(429103.1, 429259.4, by=1), y=seq(6771392, 6771729,
## by=1))
## gi <- polygrid(gr3,bor=borda3)
## kc3 <- krige.control(obj = MMv.ml)
## d.k3 = krige.conv(MMv, loc=gr3, krige=kc3)
## image(d.k3, loc=gr3, border=borda3)
## title(main = "FMA - vol/ha ", font.main=4)
```

#### 4 VALIDAÇÃO DO MODELO

```
## vc1<-xvalid(CAv, model=kc1, reestimate=T)
## plot(vc1)
## vc2<-xvalid(FSv, model=kc2, reestimate=T)
## plot(vc2)
##vc3<-xvalid(MMv, model=kc3, reestimate=T)
## plot (vc3)
```

#### 5 FUNÇÕES DOS MODELOS COMPARTILHADOS

```
## Estima os parâmetros dos modelos espaciais
```

```

## blockdiag <- function(x){
##   if(!is.list(x)) stop("x not a list")
##   n <- length(x)
##   if(n==0) return(NULL)
##   x <- lapply(x, function(y) if(length(y)) as.matrix(y) else stop("Zero-length
##component in x"))
##   d <- array(unlist(lapply(x, dim)), c(2, n))
##   rr <- d[1,]
##   cc <- d[2,]
##   rsum <- sum(rr)
##   csum <- sum(cc)
##   out <- array(0, c(rsum, csum))
##   ind <- array(0, c(4, n))
##   rcum <- cumsum(rr)
##   ccum <- cumsum(cc)
##   ind[1,-1] <- rcum[-n]
##   ind[2,] <- rcum
##   ind[3,-1] <- ccum[-n]
##   ind[4,] <- ccum
##   imat <- array(1:(rsum * csum), c(rsum, csum))
##   iuse <- apply(ind, 2, function(y, imat) imat[(y[1]+1):y[2],(y[3]+1):y[4]], imat=imat)
##   iuse <- as.vector(unlist(iuse))
##   out[iuse] <- unlist(x)
##   return(out)
## }

## parMatrix <- function(mu, sigma2, phi, tau2, n.datasets){
##   #n.datasets <- length(data)
##   D <- diag(n.datasets)
##   mu.M <- D[mu, 1:length(unique(mu)), drop=FALSE]
##   sigma2.M <- D[sigma2, 1:length(unique(sigma2)), drop=FALSE]
##   phi.M <- D[phi, 1:length(unique(phi)), drop=FALSE]
##   tau2.M <- D[tau2, 1:length(unique(tau2)), drop=FALSE]
##   return(list(mu.M, sigma2.M, phi.M, tau2.M))

```

```
## }
```

```
## Monta a matriz de variâncias e covariâncias
```

```
##montaSigma <- function(s, phi, tau, U){
```

```
## Sigma <- as.matrix(s^2 * exp(-U/phi))
```

```
## diag(Sigma) <- s^2 + t^2
```

```
## return(Sigma)
```

```
## }
```

```
## Calcula as estimativas dos parâmetros
```

```
## blockdiag <- function(x){
```

```
##   if(!is.list(x)) stop("x not a list")
```

```
##   n <- length(x)
```

```
##   if(n==0) return(NULL)
```

```
##     x <- lapply(x, function(y) if(length(y)) as.matrix(y) else stop("Zero-length  
##component in x"))
```

```
##   d <- array(unlist(lapply(x, dim)), c(2, n))
```

```
##   rr <- d[1,]
```

```
##   cc <- d[2,]
```

```
##   rsum <- sum(rr)
```

```
##   csum <- sum(cc)
```

```
##   out <- array(0, c(rsum, csum))
```

```
##   ind <- array(0, c(4, n))
```

```
##   rcum <- cumsum(rr)
```

```
##   ccum <- cumsum(cc)
```

```
##   ind[1,-1] <- rcum[-n]
```

```
##   ind[2,] <- rcum
```

```
##   ind[3,-1] <- ccum[-n]
```

```
##   ind[4,] <- ccum
```

```
##   imat <- array(1:(rsum * csum), c(rsum, csum))
```

```
##   iuse <- apply(ind, 2, function(y, imat) imat[(y[1]+1):y[2],(y[3]+1):y[4]], imat=imat)
```

```
##   iuse <- as.vector(unlist(iuse))
```

```
##   out[iuse] <- unlist(x)
```

```
##   return(out)
```

```
## }
```

```
## parMatrix <- function(mu, sigma2, phi, tau2, n.datasets){
##   #n.datasets <- length(data)
##   D <- diag(n.datasets)
##   res <- list()
##   #if(!missing(mu))
##     res$mu.M <- D[mu, 1:length(unique(mu)), drop=FALSE]
##   #if(!missing(sigma2))
##     res$sigma2.M <- D[sigma2, 1:length(unique(sigma2)), drop=FALSE]
##   #if(!missing(phi))
##     res$phi.M <- D[phi, 1:length(unique(phi)), drop=FALSE]
##   if(!missing(tau2) && !is.null(tau2))
##     res$tau2.M <- D[tau2, 1:length(unique(tau2)), drop=FALSE]
##   return(res)
## }
```

```
## nll.geo <- function(x){
##   # X é uma lista que deve conter ao menos os seguinte elementos
##   # $par : vetor de parametros (mu, sigma2, phi e tau2)
##   # $Umat : matrix de distncias entre as localizacoes dos pontos
##   # $data : dados
##   x$par[-1] <- (exp(x$par[-1]))^2
##   S <- x$par[2] * exp(-x$Umat/x$par[3])
##   diag(S) <- diag(S) + x$par[4]
##   R <- chol(S)
##   invRe <- forwardsolve(t(R), x$data - x$par[1])
##   ll <- -(length(x$data)*log(2*pi)/2 + sum(log(diag(R)))) + crossprod(invRe)/2
##   return(drop(-ll))
## }

## geoAreas <- function(mu, sigma2, phi, tau2, fix.tau2 = FALSE, data,...){
##   ## data é uma lista de objetos da classe "geodata"
##   n.areas <- length(data)
##   if(fix.tau2)
```

```

## lpars <- parMatrix(mu=mu, sigma2=sigma2, phi=phi, n.datasets=n.areas)
## else
## lpars <- parMatrix(mu=mu, sigma2=sigma2, phi=phi, tau2=tau2,
n.datasets=n.areas)
## valores iniciais "default"
## mus <- sapply(data, function(x) mean(x$data))
## vars <- sapply(data, function(x) var(x$data))
## phis <- sapply(data, function(x) min(dist(x$coords)))
## if(fix.tau2){
## s2s <- vars - tau2
## t2s <- tau2
## }
## else{
## s2s <- 0.9 * vars
## t2s <- 0.1 * vars
## }
## if((fix.tau2) && any(tau2 > vars)) stop("valores fixos de tau2 nao podem ser
##maiores que a variancia dos dados")
## allpars <- c(mus, log(sqrt(s2s)), log(sqrt(phis)))
## if(!fix.tau2) allpars <- c(allpars, log(sqrt(t2s)))

## lpars.std <- lapply(lpars, function(x) t(x)/colSums(x))
## ## na dimensao do vetor de parametros a ser estimado:
## inis <- drop(blockdiag(lpars.std) %*% allpars)
## ## para todas as áreas
## inisM <- matrix(blockdiag(lpars) %*% inis, nrow=n.areas)

## distancias em cada dado (sao fixas na maximizacao)
## for(i in 1: n.areas)
## data[[i]]$Umat <- as.matrix(dist(data[[i]]$coords, diag=TRUE, upper=TRUE))
## fc para calcular neg log-lik para uma area
## nll <- function(par, lpars, n.areas, data){
## parM <- matrix(blockdiag(lpars) %*% par, nrow=n.areas)
## if(fix.tau2) parM <- cbind(parM, log(sqrt(t2s)))

```



```

##   for(i in 1:n.areas) data[[i]]$par <- parM[i,]
##   nll.areas <- sapply(data, nll.geo)
##   return(sum(nll.areas))
## }

## res <- optim(inis, nll, lpars=lpars, n.areas=n.areas, data=data, ...)
## parM <- matrix(blockdiag(lpars) %*% res$par, nrow=n.areas)
## parM[,-1] <- (exp(parM[,-1]))^2
## if(fix.tau2) parM <- cbind(parM, t2s)
##   dimnames(parM) <- list(paste("area", 1:length(data), sep="_"),
c("mu","sigma2","phi","tau2"))
## res <- c(list(parM = parM, estpars = res$par, npars = length(res$par),
## loglik = -res$value, nobs = sum(sapply(data, function(x) ##nrow(x$coords))),
method.lik = "ML"), res[-c(1:2)])
##   oldClass(res) <- c("multiGRF")
##   return(res)
## }

## logLik.multiGRF <- function (object, ...)
## {
##   val <- object$loglik
##   attr(val, "df") <- object$npars
##   attr(val, "nobs") <- object$nobs
##   attr(val, "method") <- object$method.lik
##   class(val) <- "logLik"
##   return(val)
## }

## dd <- list(list(coords=matrix(runif(12), nc=2), data=round(rnorm(6, m=100, sd=5))),
##   list(coords=matrix(runif(8), nc=2), data=round(rnorm(4, m=130, sd=3))),
##   list(coords=matrix(runif(14), nc=2), data=round(rnorm(7, m=80, sd=2))))
## debug(geoAreas)

## Fim da programação.

```